

Mohammad Salem

# Introduction To Basic Concepts Of Medical Genetics

Molecular genetics & Pathogenetics



**LAMBERT**  
Academic Publishing

**Mohammad Salem**

# **Introduction To Basic Concepts Of Medical Genetics**

**Molecular genetics & Pathogenetics**

**LAP LAMBERT Academic Publishing**

# Contents

## Part 1

### **Basic concepts of molecular medical genetics 3**

Scope of medical genetics	3
Spectrum of medical genetics	3
Genetics and life	4
Structure of the genetic material	5
Structural differences between DNA and RNA	8
Structural organization of the human genome	10
Functional organization of the human genome	13
Pseudogenes	13
Pyknons	14
Transposons	15
Telomeres	16
Structure of human genes	17
Functions of human genes	17
Stages of gene function	18
Gene activation	18
Transcription	19
Post-transcription modifications	19
Translation	21
Post-translation modifications	22
Post-translation trafficking of proteins	23

## Part 2

### **Basic concepts of medical pathogenetics 24**

Pathogenetics	24
Mutagens	24
Classification of mutagens	24
Factors affecting actions of mutagens	26
Mutation	26
Types of mutation	26
1. Induced versus spontaneous mutation	27
2. Nuclear versus mitochondrial mutation	27
3. Somatic versus germinal mutation	28
4. Pathological versus non-pathological mutation	28
5. Static versus dynamic mutation	29
6. Persistent versus reversible mutation	30
7. Point, Small, Gross, Genomic mutation	30
a. Point mutation	30
Molecular mechanisms of point mutation	31
Types of point mutation	32
b. Small mutations	33

c. Gross mutations	34
Structural chromosomal aberrations	36
Numerical chromosomal aberrations	39
Molecular mechanisms of chromosomal aberrations	39
d. Genomic mutations	40
Numerical genomic mutations	40
Structural genomic mutations	41
Functional genomic mutations	42
8. Base, Sugar, Phosphate group mutation	42
Pathogenesis of genetic diseases	43
The concept of protein domain	43
The concept of protein mutation	44
Pathogenetic mechanisms	45
Anti-mutation mechanisms of the human genome and human proteome	46
Structural organization of the human genome	48
Structural features of DNA	48
Degeneracy of the genetic code	48
Nuclear localization of DNA	48
DNA-associated proteins	49
Replication proofreading system	49
Genetic repair systems	49
Nuclear DNA repair	49
RNA repair/editing system	51
Mitochondrial DNA (mtDNA) repair	51
Protein repair systems	51
Silencing of transposon activity during development	52
Antioxidant enzyme systems	53
Apoptosis of heavily, lethally mutated cells	53
Melatonin	53
References	54

## **Part 1: Basic concepts of molecular medical genetics**

### **1. Scope of medical genetics**

Medical genetics is a branch of human genetics. It is concerned with, and confined to, the study of the effects of the genetic material on health and disease states of human beings. Within this context, it comprises the study of the structure and functions of genes in normal and in pathological conditions, the study of causes and mechanisms of pathogenesis of genetic disorders, characterization of the different types of these disorders and delineating the variable modes of their inheritance.

Diagnosis of genetic diseases, prophylaxis against their occurrence and/or their complications and invention of effective therapeutic measures aiming at obviating or alleviating their consequences on affected subjects comprise major aspect of the wide spectrum of medical genetics. In addition, the merging and integration of available databases of **basic genetics**, **clinical genetics** and **diagnostic genetics** allows their application in many different ways for specific purposes. Of utmost importance among these applications is the use of available and invented treatment modalities in management of patients with genetic diseases. Conventional drug therapy, nutritional control and organ transplantation for many of these diseases are just few examples within this context. The use of conventional treatment modalities as well as of invented approaches, like stem cell therapy and antisense oligonucleotides, in management of genetic disorders are the main features of the field of **therapeutic genetics**. Nevertheless, the integration and application of available genetic knowledge is progressing in a steady way to include other important health aspects. For instance, anticipation of possible health hazards based on the genotype of the individual or the family history allows for implementing specific measures to avoid and/or alleviate possible consequences. This approach in management of genetic diseases constitutes the cornerstone of **prophylactic genetics**. Similarly, application of genetic databases for offering genetic counseling to patients and to their concerned family members, in newborn screening, in forensic medicine and in fetal therapy, as well as in many other practical aspects, constitute some of the main achievements of **applied genetics**.

### **2. Spectrum of medical genetics**

1. Basic genetics (Table 1)
2. Clinical genetics (Table 2)
3. Diagnostic genetics (Table 3)
4. Therapeutic genetics (Table 4)
5. Prophylactic genetics (Table 5)
6. Applied genetics (Table 6).

**Table 1: Branches of basic genetics**

1. Molecular genetics	10. Formal genetics
2. Pathogenetics	11. Population genetics
3. Cytogenetics	12. Evolutionary genetics
4. Biochemical genetics	13. Comparative genetics
5. Physiological genetics	14. Experimental genetics
6. Pharmacogenetics	15. Genomics
7. Oncogenetics	16. Transcriptomics
8. Immunogenetics	17. Proteomics
9. Developmental genetics	

**Table 2: Branches of clinical genetics**

1. Chromosomal aberrations	12. Genetic diseases of the muscular system
2. Congenital malformations	13. Genetic diseases of the skeletal system
3. Inborn errors of metabolism	14. Genetic diseases of the blood
4. Mitochondrial disorders	15. Genetic diseases Of the immune system
5. Genetic systemic syndromes	16. Genetic diseases of male reproductive system
6. Genetic diseases of the nervous system	17. Genetic diseases of female reproductive system
7. Genetic diseases of the endocrinal system	18. Genetic diseases of the eye and vision
8. Genetic diseases of the cardiovascular system	19. Genetic diseases of the ear and hearing
9. Genetic diseases of the respiratory system	20. Genetic diseases of the skin
10. Genetic diseases of the GIT system	21- Genetic basis of psychiatric disorders
11. Genetic diseases of the urinary system	

**Table 3: Branches of diagnostic genetics**

1. Molecular diagnostic techniques: e.g. DNA microarray
2. Cytogenetic diagnostic techniques: e.g. high resolution banding
3. Biochemical diagnostic techniques: e.g. enzyme assay
4. Conventional diagnostic techniques: e.g. radio-imaging, electro-physiological studies

**Table 4: Branches of therapeutic genetics**

1. Pharmacological therapy	5. Surgical intervention
2. Nutritional therapy	6. Stem cell therapy
3. Enzyme replacement therapy	7. Genetic therapy
4. Organ transplantation therapy	8. Conventional therapy

**Table 5: Branches of prophylactic genetics**

1. Premarital prophylaxis: e.g. premarital counseling.
2. Pre-conception prophylaxis: e.g. pre-implantation diagnosis.
3. Prenatal prophylaxis: e.g. congenital adrenal hyperplasia.
4. Pre-symptomatic prophylaxis: e.g. hemochatosis.

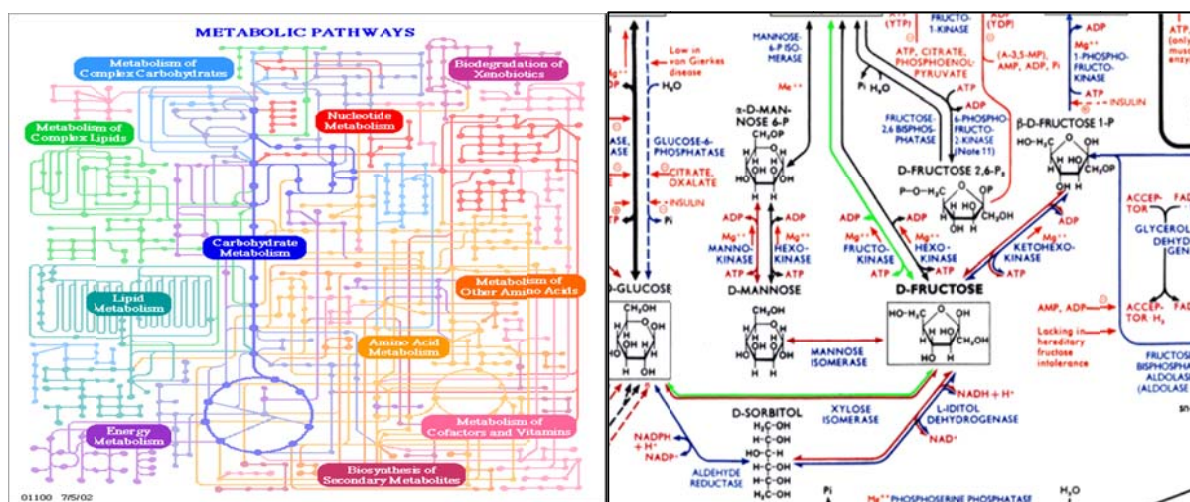
**Table 6: Branches of applied genetics**

1. Genetic counseling
2. Genetic screening: e.g. newborn screening, carrier screening
3. Pre-implantation diagnosis: single gene disorders
4. Prenatal diagnosis
5. Fetal Therapy
6. Forensic Genetics
7. Genetic Engineering
8. Eugenics

### 3. Genetics and life

All life activities in living cells, whether conducted on molecular level like electron transfer in oxidative phosphorylation for ATP production, on cellular level like cell division, on tissue level like muscle contraction or on whole organ level like hearing for instance, are mediated via a very large number of inter-related **metabolic networks**. A metabolic network is defined as a cascade of controlled biochemical reactions and biophysical alterations that transform one, or more, substrate to one, or more, product. Each metabolic network consists of a very large number, sometimes thousands, of proteins, mostly enzymes, and other non-protein factors all acting co-operatively in sequence to perform specific biochemical and physiological functions. In human cells, nearly 4100 (four thousand and one hundred) of these networks have been delineated [Hatzimanikatis V. et al, 2004] (Figure1).

Proteins and enzymes which are the major mediators and determinants of all metabolic networks in living cells are synthesized under direct and strict regulation of the genetic material. The structural genes, which are the major component of the genetic material, are primarily concerned with controlling and regulating the synthesis of proteins, which in turn control and regulate all life activities in living cells.

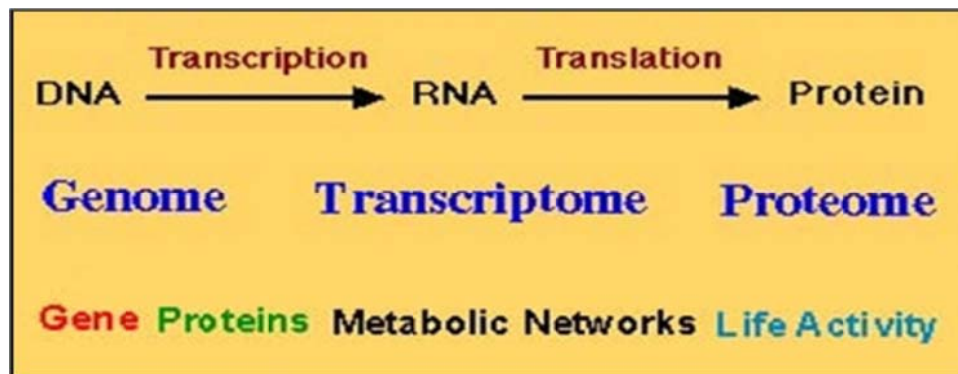


**Figure 1: Metabolic networks**

[www.genome.ad.jp/kegg/kegg.html](http://www.genome.ad.jp/kegg/kegg.html)

[www.expasy.org/tools/pathways](http://www.expasy.org/tools/pathways)

This intimate relationship between the genetic material and life activities is represented in **(Figure 2)** which summarizes this relationship and clarifies the central dogma of molecular biology which states that while the sum total of the genetic material of the living cell, collectively referred to as the **genome**, controls and encompass the whole spectrum of life processes in living cells, the sum total of the proteins produced and synthesized in the cell under strict control of the genome, collectively referred to as the **proteome**, are the actual and direct mediators of these life processes.



**Figure 2: Relation between the genetic material and life activities**

#### **4. Structure of the genetic material**

The building components of the genetic material in all living creatures are the **nucleic acids**. There are two main categories of nucleic acids: DNA or Deoxyribo-Nucleic Acid, **DNA**, and RNA or Ribo-Nucleic Acid (**Figure 3**). With the exception of RNA viruses which have their genome composed solely of RNA, all living creatures have DNA as their sole genetic material in addition to RNA as well.

Nucleic acids are long linear, strand-shaped unbranched hetero-polymers, composed of large number of similar monomers: the **nucleotides**, which are the building blocks of the nucleic acids. Each nucleotide is composed of an inorganic phosphate group attached to a 5-carbon atom sugar, a ribose sugar in RNA and a 2-deoxyribose sugar in DNA, to which is attached a nitrogenous base (**Figure 4**).

Five different bases participate in formation of five different nucleotides that build up the nucleic acids. The bases are either purine bases: adenine (**A**) and guanine (**G**), or pyrimidine bases: cytosine (**C**), thymine (**T**), and uracil (**U**). The nucleotides are usually referred to by the type of base they contain, hence we have (T), (C), (G), (A) and (U) nucleotides. The first four nucleotides are found exclusively in DNA, and Uracil replaces Thymine in RNA (**Figure 5**).



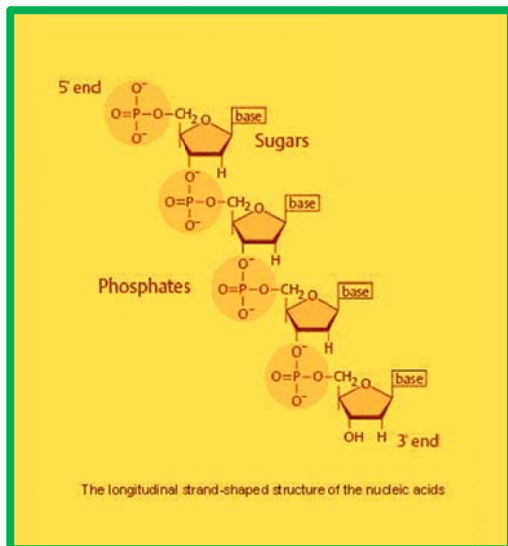


Figure 3

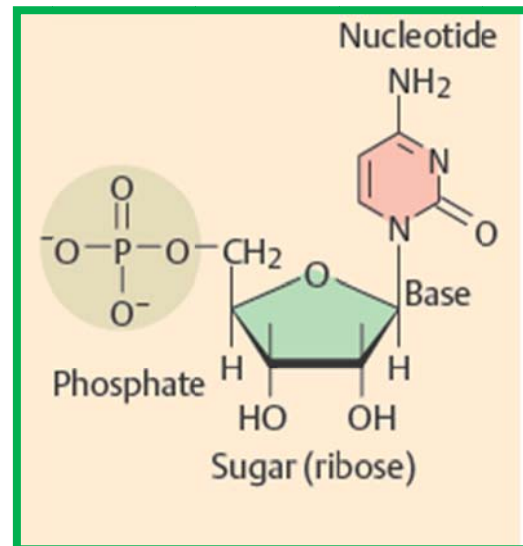


Figure 4

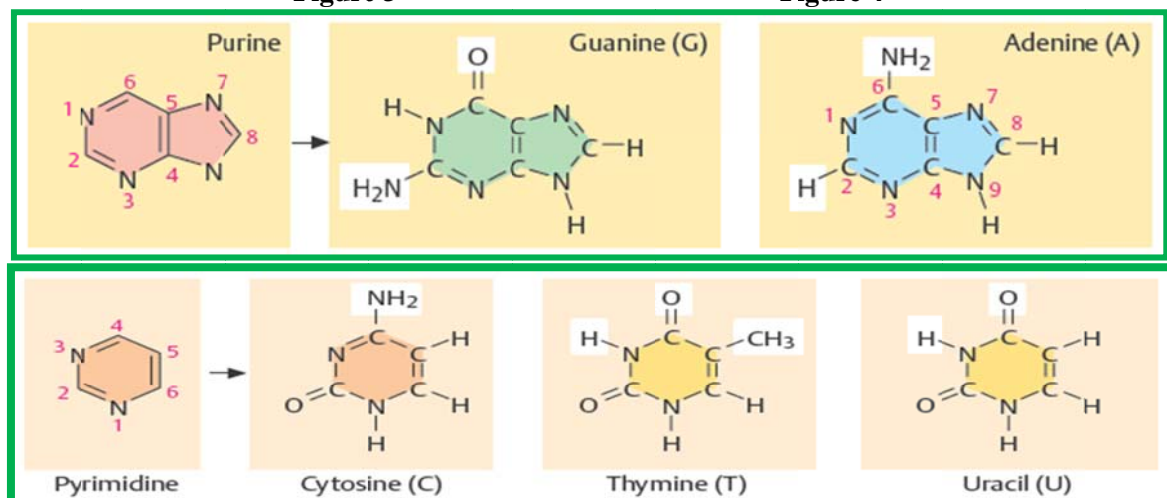


Figure 5: The nucleotides of nucleic acids

The longitudinal strand-shaped structure of the nucleic acids is maintained by the side-by-side attachment of the nucleotides, with the phosphate group of one nucleotide being attached to the ribose sugar of the next nucleotide. DNA occurs naturally as a double stranded structure composed of two complementary strands attached together by the hydrogen bonds of the nitrogenous bases of each two opposing nucleotides. In addition to stabilizing and strengthening the DNA structure, the second or complementary strand of DNA acts as a template for synthesis of identical parts of the other strand in case of loss or damage. The double stranded configuration of DNA offers more stability to DNA via protecting the bases from being damaged by reactive radicals generated during metabolic activities of the cell. If damage to the gene occurs, the other complementary strand is used as an information or code template to repair the gene and replace defective parts of it by a specific DNA repair system.

The concept of **base complementarity**, which is a key feature of DNA structure, integrity and stability, entails that the four nucleotides, referred to simply as bases, occur along the opposing strands of the DNA in a specific order so that whenever

adenine exists on one strand it always combines with thymine on the opposing strand (**A-T complementarity**) and that whenever cytosine exists on one strand it always combines with guanine on the other strand (**C-G complementarity**). The same rules of base complementarity hold true for double-stranded species of RNA (**Figure 6**).

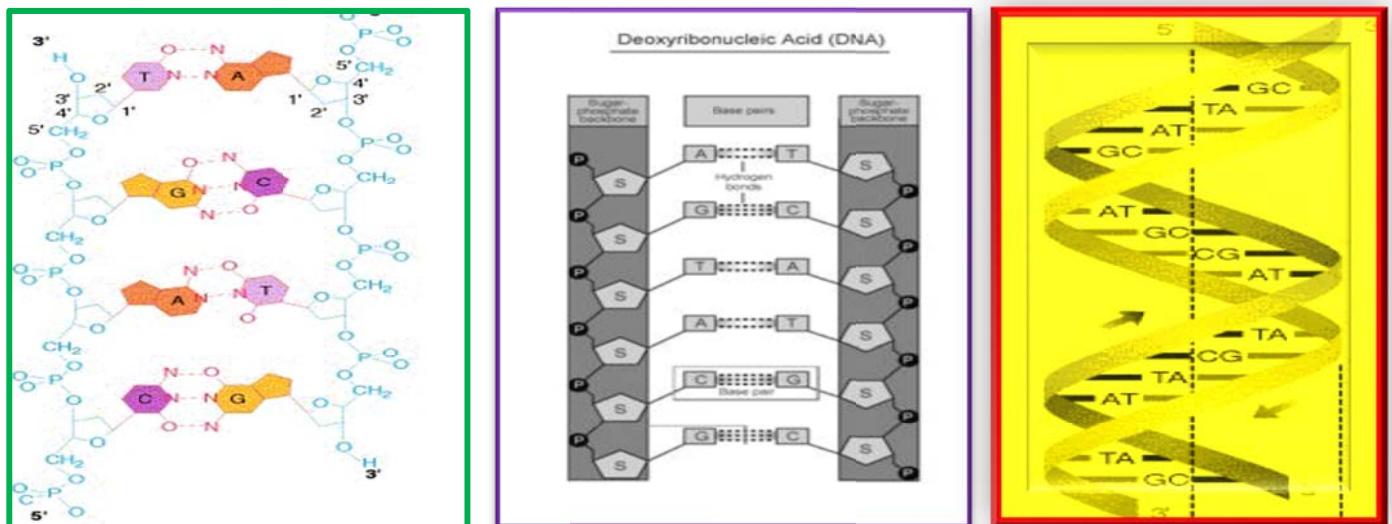


Figure 6: The concept of base complementarity

During cell division, preservation of base complementarity between the original and the newly synthesized strands, during **replication of DNA**, is mandatory since it guarantees the synthesis and production of identical DNA double stranded molecules, each consisting of half original strand from the parent molecule and half new strand. This mode of DNA replication is referred to as **semiconservative replication** because the new DNA contains only half the content of the original DNA molecule. This peculiar way of DNA replication allows for duplication of the nuclear genome of the mother cell during cell division and its equal distribution between the resulting daughter cells. More crucial, however, is that it maintains **identity of the nuclear genome** of all cells descendent from the mother cell, as all of them contain identical nuclear genomes, all sharing half the original genome of the parent cell (**Figure 7**).

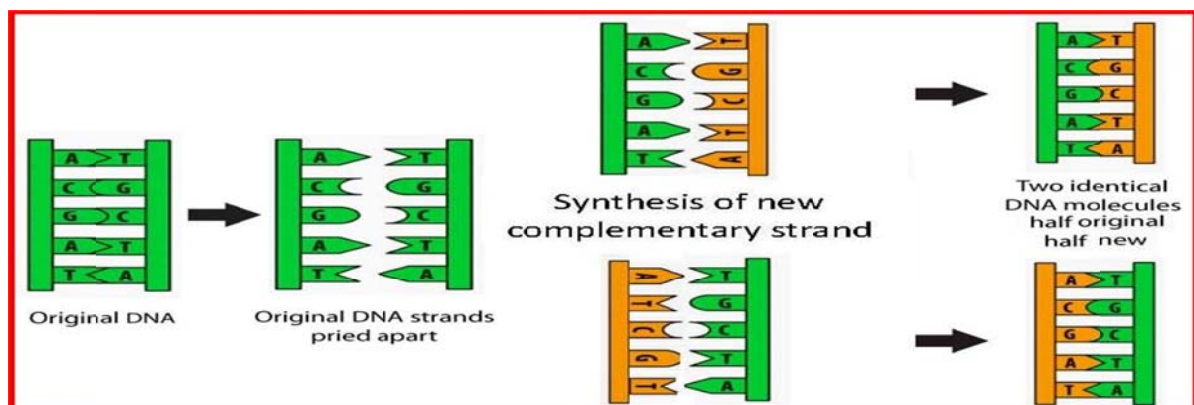


Figure 7: Semiconservative replication of DNA

## 5. Structural differences between DNA and RNA

Nuclear DNA normally exists as a linear unbranched double stranded helix composed of two tightly coupled strands. Each strand is a linear hetero-polymer of the four types of nucleotides referred to (A,T,G and C). Each nucleotide is attached to the two adjacent nucleotides via phospho-di-ester bond between its phosphate group and the sugar of the adjacent nucleotide, i.e. the phosphate of one nucleotide is attached to the carbon no. 5 of the sugar of its adjacent nucleotide, whose phosphate in turn is linked to the carbon no. 3 of the sugar of the next nucleotide, and so on, so that a phosphate-sugar-phosphate-sugar, etc. linear strand, forms the backbone of the DNA, with the nitrogenous base attached to the sugar, through a glycosidic linkage, projecting at nearly right angles to the level of this backbone. Accordingly, the direction of phosphate-sugar bonding on one strand goes in a 5' to 3' direction, whereas the direction of phosphate-sugar bonding on the other complementary strand goes in the opposite 3' to 5' direction. This particular arrangement, referred to as **DNA strand polarity**, is important because gene transcription always proceeds in a 5' to 3' direction (**Figure 8**).

The two strands of DNA are tightly attached to each other by hydrogen bonds between the nitrogenous bases of each opposing nucleotide pair at the same position of the DNA strands. Normally, DNA does not exist *per se*. It is wrapped by, and associated with, a heavy coat of **DNA-binding proteins** of two classes: the histones and the non-histone proteins. Histones are small proteins with a very high proportion of positively charged amino acids (lysine and arginine), this characteristic helps tight binding to the highly negatively charged phosphate of the DNA. This tight bonding plays a crucial role in maintaining DNA-Histone association necessary for support, protection, and regulation of DNA structural and functional integrity. There are five types of histones which participate in formation of the DNA-associated protein coat: small, highly conserved nucleosomal histones (H2A, H2B, H3, and H4) and the H1 histones. The nucleosomal histones form a specific disk-shaped complex of eight proteins containing two copies of each of the four nucleosomal histones, known as the **histone octamer**. This histone octamer forms a protein core around which the double-stranded DNA helix is wound twice. This characteristic DNA-Histone complex is known as the **nucleosome**. The histone H1 molecules, of which there are about six closely related subtypes in eukaryotic cells, are thought to be responsible for associating or binding nucleosomes together thus imparting to the DNA-Histone complex its fibrillar or linear strand conformation.

RNA, or ribonucleic acid, has the same basic design of a sugar-phosphate backbone with a nitrogenous base linked to the sugar. With few exceptions, RNA exists as a single stranded structure. However, RNA differs from DNA in many aspects (**Figure 9**).

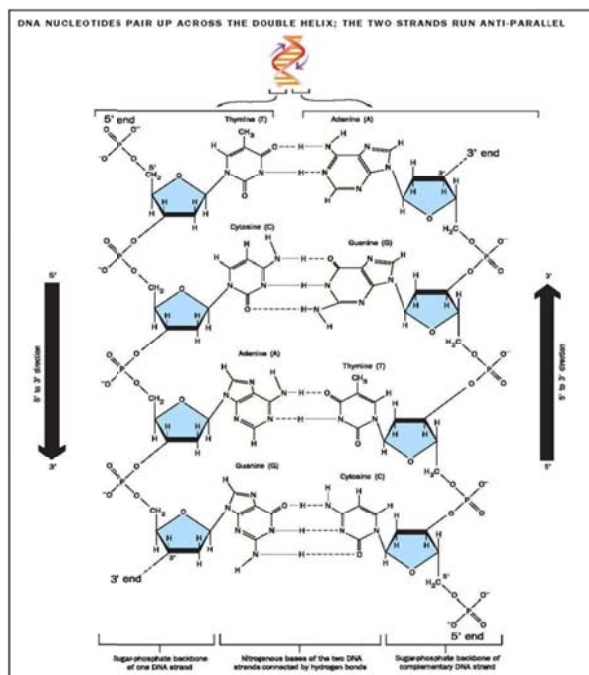


Figure 8: Strand polarity of DNA

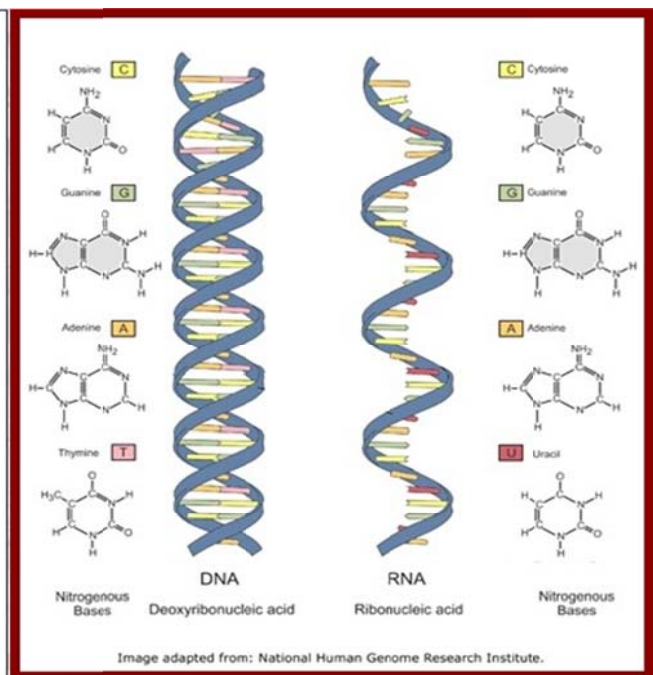


Figure 9: Structural differences between DNA and RNA

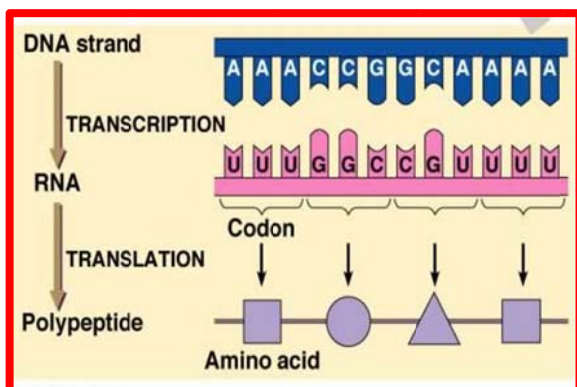
1. DNA normally exists in the nucleus as very long linear double stranded helix and in the mitochondria as circular double stranded tiny molecule, whereas RNA normally exists in the nucleus, mitochondria and cytoplasm as very short linear, single and double stranded structures.
2. The sugar in DNA is 2-deoxy-ribose lacking a oxygen atom at carbon 2, whereas it is ribose in RNA.
3. DNA is composed of A,G,C and T, whereas RNA is composed of A,G,C and U.
4. The single stranded structure of RNA confers upon it the ability to form complex three-dimensional molecular shapes that cannot be formed by double-stranded DNA. It can fold into characteristic secondary and tertiary structures that account for its diverse functional activities.
5. Whereas DNA normally exists as single functional type that performs both replication and transcription, there are at least **four different functional subtypes of RNA**: **mRNA** or **messenger RNA**, which is the transcript of the gene having the genetic information coded by both exons and introns and responsible for carrying these information for protein synthesis to the site of synthesis in the cytoplasm, **rRNA** or **ribosomal RNA** responsible for decoding or translation of the genetic information in mRNA to specific amino acid sequence in the synthesized protein, **tRNA** or **transfer RNA** responsible for participation in decoding mRNA and getting the required amino acids for protein synthesis from the metabolic pool of the cell and transporting them to the site of translation and synthesis in the ribosomes and **miRNA** or **micro RNA** which are **non-coding** species of RNA, i.e. they are not transcripts of genes and are not translated to proteins.

Micro RNA, sometimes referred to as small RNA, comprise many subtypes and play critical roles in many functional aspects of the genetic material, including regulation

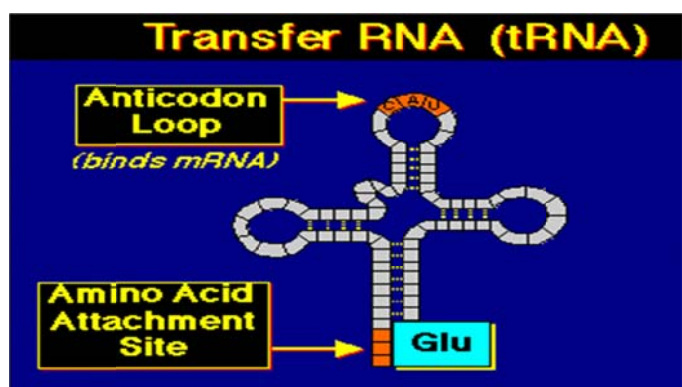


of gene function. Among the important subtypes of miRNA is **guide RNA** or **gRNA**, involved in RNA editing, **small cytoplasmic RNA** or **scRNA**, that participates in protein trafficking and targeting in the cell, **ribozymes** or **catalytic RNA** molecules with specific enzymatic activities and **piRNA** or **piwiRNA** species which play a fundamental role in stabilizing the genome during development due to their roles in suppressing and silencing transposon activities (**Figure 10**).

Small **piRNAs** are composed of (26–31) nucleotides with 5' monophosphate and peculiar 3' modification (2'-O-methylation) that has been suggested to increase stability of the molecule, probably via reducing its destruction by active oxidant radicals. Small piRNA classes do not have secondary structure, they lack sequence conservation and comprise many subtypes that are found in the nucleus and the cytoplasm of germ cells, particularly in male germ cell lines. They constitute the largest proportion of small non-coding RNA species in both vertebrates and invertebrates cells and it is estimated that mammalian cells contain many hundreds of thousands of different piRNA species. They exert important roles in epigenetic and post-transcriptional gene silencing of retro-transposons and other genetic elements in germ line cells.



mRNA (Messenger RNA)



tRNA (Transfer RNA)

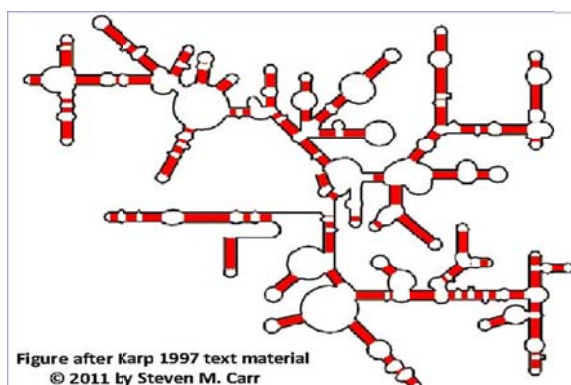
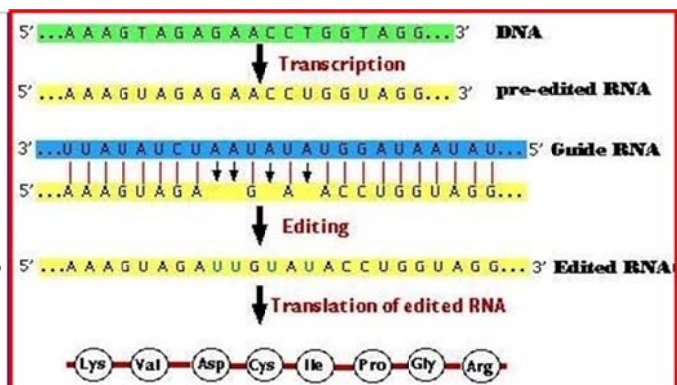
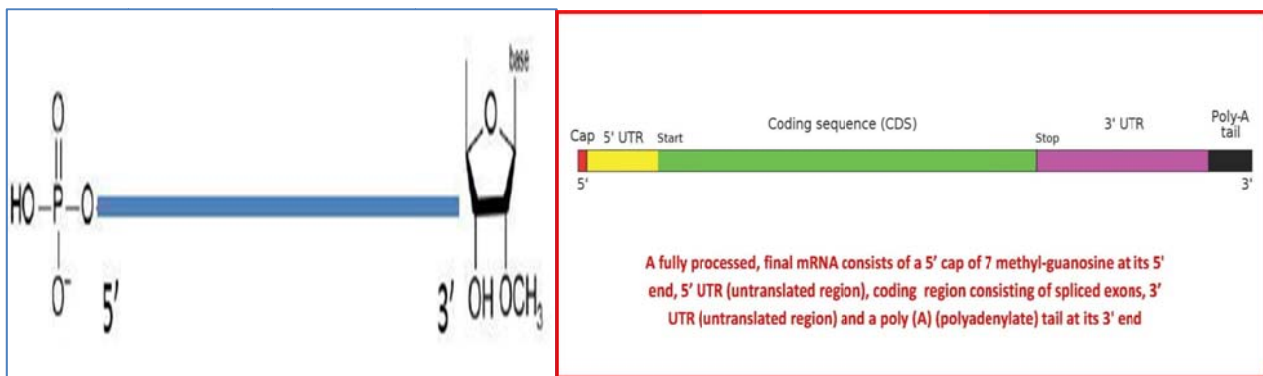


Figure after Karp 1997 text material  
© 2011 by Steven M. Carr



rRNA (Ribosomal RNA) gRNA (Guide RNA) [www.users.rcn.com/~jkimball.ma.ultranet/BiologyPages](http://www.users.rcn.com/~jkimball.ma.ultranet/BiologyPages)  
Figure after Karp 1997 text material, © 2011 by Steven M. Carr



piRNA or Piwi-interacting RNA

Structure of mature mRNA

Figure 10: Functional types of small or microRNA

The sum total of all RNA species in the cell is referred to as the **transcriptome**. However, not all types of RNA are directly involved in protein synthesis in the cell. RNA species that carry genetic information necessary for protein synthesis are known as **coding RNA** and include mainly mRNA. Other RNA species that do not code for protein synthesis are named **non-coding RNA** species. They include several subtypes of RNA like miRNA and piRNA.

## 6. Structural organization of the human genome

The **genome**, or the sum total of the genetic material in the cell, consists of genes in addition to other non-genic portions of DNA, sometimes referred to as intervening sequences or inter-genic DNA, and gene-related sequences. All the genes as well as all other non-genic parts of the genome, however, are composed of the same four nucleotides that constitute the building blocks of the DNA. They differ only in the number and specific arrangement of their nucleotide content. Each species has its own specific genome that differs from the genome of any other species as regards the number of genes, their cellular distribution and the size of the genome itself, among many other inter-species differences (**Figure 11**).

### 6.1. Nuclear genome

In human cells, the human genome is unequally distributed into a major part, constituting more than 99.999 % of its size, organized in the form of long strands, open-ended **chromosomes** contained in the nucleus and referred to as the **nuclear genome (Figure 12)** which comprises between 20000–25000 genes distributed over the chromosomes. Each chromosome consists of a very long double stranded molecule of DNA wrapped with a heavy coat of basic proteins composed mainly of histones and protamines. These DNA-associated proteins offer support and protection to the DNA and play a critical role in regulating many aspects of gene functions.

The nuclear genome in each human germ cell, ovum and sperm, is organized into a set of 23 separate chromosomes known as the **haploid genome** which represents the **unit genome** of humans. Upon fertilization, both haploid genomes of the sperm and the ovum constitute a **diploid genome** consisting of their 46 chromosomes that

characterizes the nuclear genome of the zygote as well as of all somatic cells descendant from it. With very few exceptions, the sperm does not contain mitochondria. Nearly all mitochondria, and hence the mitochondrial genome, present in the zygote and in all body cells are descendant from the mitochondria present in the ovum at fertilization.

Nuclear genes are arranged in a linear sequence on chromosomes. The estimated 20000 - 25000 genes that comprise the nuclear genome constitute, and are distributed on, the 46 chromosomes in the nucleus. The larger and longer chromosomes have far more numbers of genes than the smaller and shorter chromosomes. Because genes constitute only a small proportion of the whole genome, they are separated by long inter-genic parts of base sequences of the DNA that comprise most of the non-genic or gene-related components of the genome. These include: pseudo-genes, pyknons, transposons and telomeres in addition to an exceedingly large numbers of long and short repetitive and non-repetitive interspersed elements, among many other components of, yet, undefined function(s) [Mingyao Li et al, 2011].

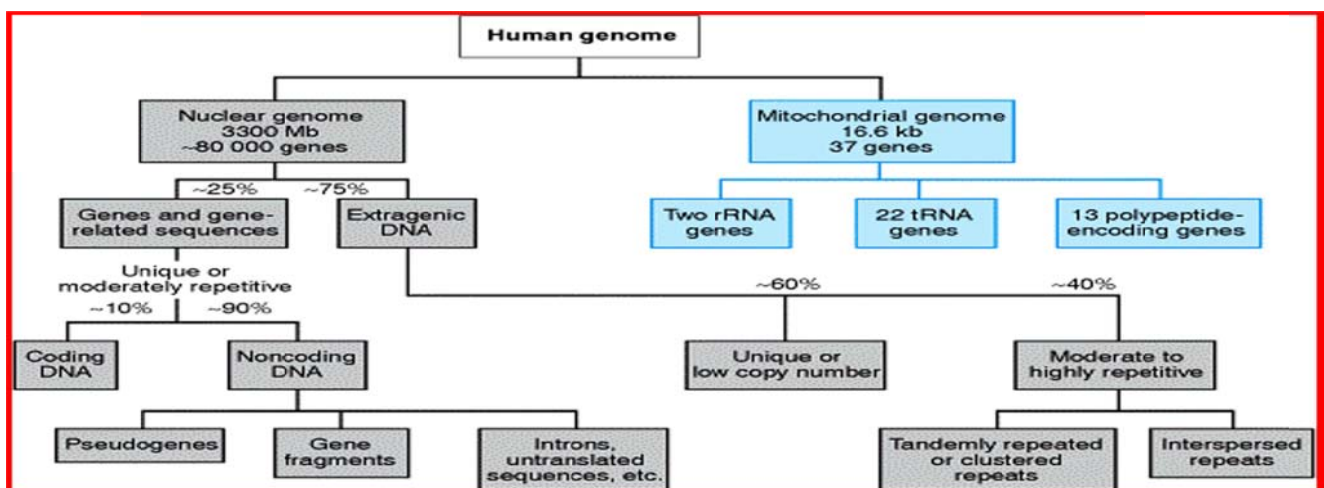


Figure 11: Structural organization of the human genome

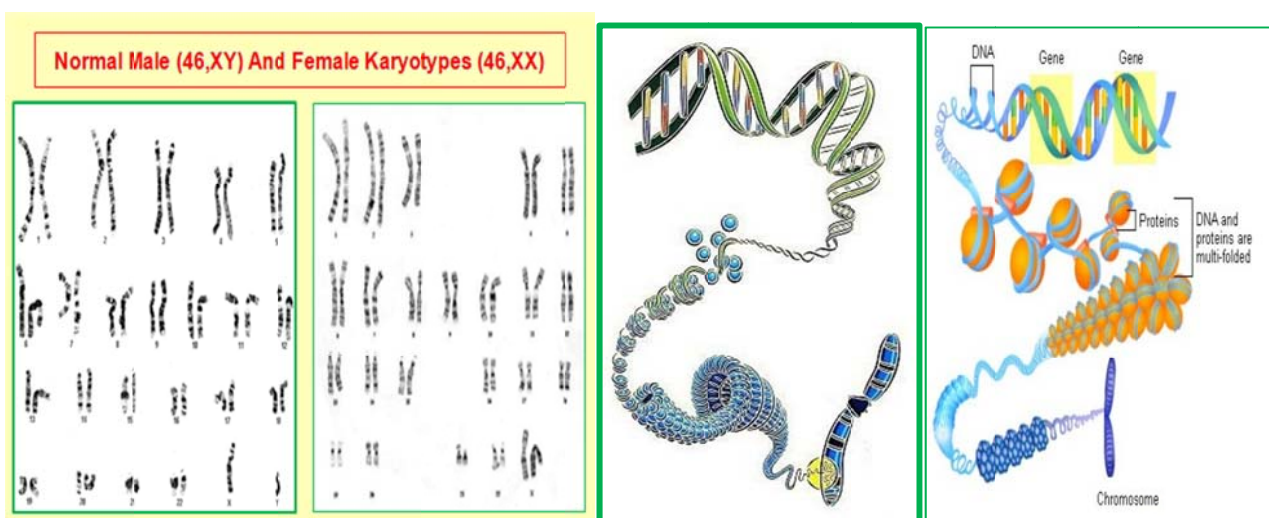
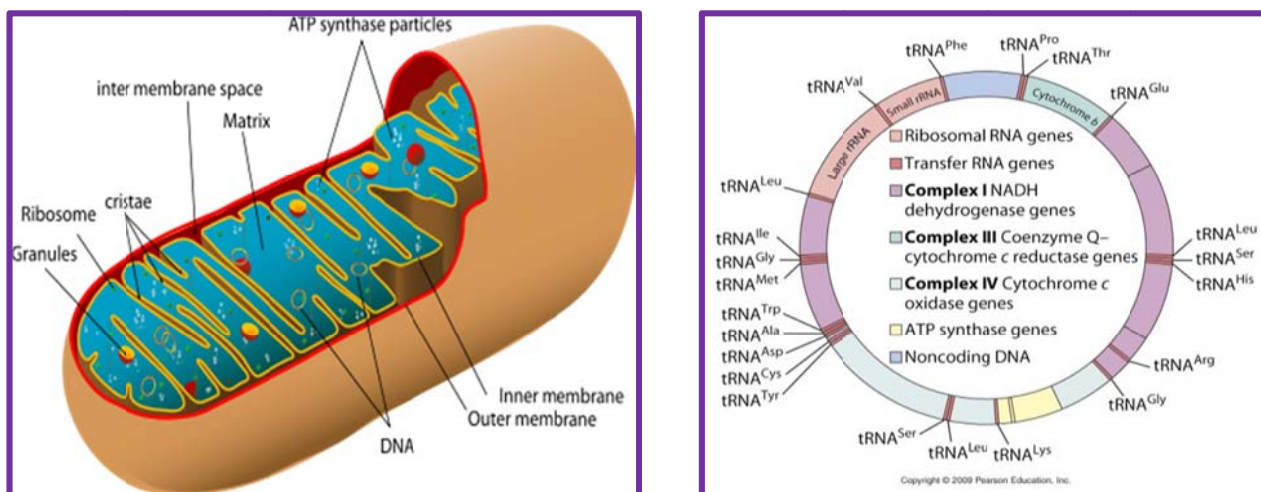


Figure 12: Structural organization of the human nuclear genome

## 6.2. Mitochondrial genome

The remaining tiny part of the human genome exists in the form of varying numbers, tens to thousands, of very small closed circular double stranded structures present inside the mitochondria and is referred to as the mitochondrial genome (**mtDNA**). Each molecule of the **mitochondrial genome** consists exclusively of 37 genes (**Figure 13**). Though it constitutes a very tiny fraction of the whole genome, mtDNA is indispensable for life because it codes for proteins that mediate ATP production in the cell in addition to many other important functions like apoptosis and many other vital metabolic activities like lipid oxidation and steroid biosynthesis.

The number of mitochondria and the number of mtDNA molecules in each mitochondrion varies according to the metabolic activities of the cell. The most active and energy-demanding cells, like neurons, heart muscles, the retina, skeletal muscles, endocrine glands, kidney cells and liver cells have the largest numbers of mitochondria within their cytoplasm and the largest numbers of mtDNA molecules in each mitochondrion as well.



**Figure 13: Structural organization of the human mitochondrial genome**

The peculiar features of mitochondrial genetics and mtDNA make it a plausible target for gene therapy trials. It also has many obvious advantages over classic gene therapy trials aiming at repair or replacement of mutated nuclear genes. The reasons for this assumption can be summarized in the following notions:

1. The presence of mtDNA in the cytoplasm away from the nucleus obviates the known hazards associated with classic gene therapy trials directed towards nuclear genes, like uncontrolled and random insertional mutagenesis, which results in widespread damage to the nuclear genome.
2. The autonomous nature of mitochondrial replication allows for higher, more persistent production of the desired gene product due to the presence of many,



hundreds to thousands, mitochondria in each cell and presence of multiple copies of mtDNA within each mitochondrion.

3. If viral vectors are used for DNA delivery, viral genomes do not replicate within the mitochondrion due to absence of other genes necessary for their replication, thus almost all of the deleterious consequences of viral impact on host DNA are avoided.

4. The presence of large copy number of mtDNA within each mitochondrion permits a safer margin in case of failure of gene delivery or damage to the mitochondrial genome since no genetic deleterious effects happen in living cells unless nearly 80% of its mitochondrial DNA content becomes mutated or malfunctioning.

5. Use of non-viral vectors, e.g. DQAsomes, gene delivery system made from Dequalinium [Volkmar Weissig et al, 1998], for DNA delivery to mitochondria, avoids all risks known to be associated with use of viral vectors for gene delivery.

### 6.3. Cytoplasmic membrane associated DNA

A tiny portion of DNA exists in the cytoplasm of somatic cells attached to the inner side of the cytoplasmic membrane and is known as **cm-DNA** or **cytoplasmic membrane associated DNA**. It has physical and chemical properties different from both chromosomal and mitochondrial DNAs and represents a portion of the heterochromatin of the centromeric and peri-centromeric regions of chromosomes that exited to the cytoplasm. cm-DNA is transcribed separately in the cytoplasm by a specific RNA polymerase different from that used for nuclear DNA transcription. The potential functions of cmDNA are largely undefined. However, many putative roles have been assigned to it including mediation of cellular activities, e.g. control of signal transduction in the cytoplasm and induction or regulation of apoptosis. Inappropriate over-transcription of cmDNA might result in disturbance of the intricate balance between oncogenes and tumor suppressor genes and has been implicated in development of some malignancies, like breast cancer [Cheng J. et al, 2012].

## 7. Functional organization of the human genome

### 7.1. Functional genes

Though all portions of the genome consist of the same four nucleotides that compose the nucleic acids (G-C-A-T in DNA) and (G-C-A-U in RNA), a highly remarkable sophisticated degree of functional organization has been attained by this simple basic structure. The functional characterization of any part of the genome is attributed, primarily, to the number and the specific arrangement of these four bases that compose it. Functional organization of the genome might be attained by many different ways according to the parameter of classification (**Table 7**). The genome is broadly organized into **functioning genes** and apparently non-functioning inter-genic or gene-related segments. Functional genes are categorized into three main categories: **structural genes** which are directly involved in, and responsible for, protein synthesis via transcription and translation, **regulatory genes** which regulate

the functions of structural genes and **master genes** which are responsible for controlling and monitoring vital cellular processes without which the cell can't maintain its integrity and stability. These processes include cell division, cell differentiation, DNA repair, apoptosis, transport across cell channels, inter-cellular adhesion and contact and many others.

## 7.2. Pseudogenes

Pseudogenes are DNA sequences that structurally resemble functional genes. There are two types of pseudogenes known as **processed** and **unprocessed** pseudogenes. Processed genes are found on different chromosomes, they lack introns and certain regulatory elements, often terminate in adenine series, and are flanked by direct repeats. They may be complete or incomplete copies of genes or mixtures of several genes. They are believed to have occurred through transcription of the original gene to mRNA followed by posttranscriptional removal of introns and forming back DNA through a reverse transcription process. Unprocessed pseudogenes, having their original introns and associated regulatory elements, usually exist as clusters of similar structural sequences on the same chromosome. Their active expression is usually suppressed by one or more type of point or small mutation affecting its promotor area, including deletion, insertion or change to stop or termination mutation. Unprocessed pseudogenes are believed to have arisen by duplication of their parent gene [Balakirev and Francisco, 2003]. The persistence of pseudogenes as conserved structural components of the genome of human and too many other organisms suggests an important, still undefined, role(s) played by pseudogenes in genetic processes justifying their existence. Allied to the role played by the complementary strand of DNA as a reserve or a substitute for the original strand in case of failure to repair damage to the original strand, pseudogenes may have similar role in cases of catastrophic unreparable damage of the original genes. They might function as standby genes ready for repair and/or reactivation to undertake the functions of damaged genes. Also, they may have quantitative enhancing roles during early embryonic growth and differentiation where the needs for gene products are particularly most demanded by fast growing and dividing cells during this stage of development. Assigning no functions to pseudogenes seems nonsensical in view of their structural conservation, their common sharing between unrelated different organism and their exceedingly large number, about 19000 genes, in the human genome [Gibson, 1994] (Figures 14-A & 14-B).

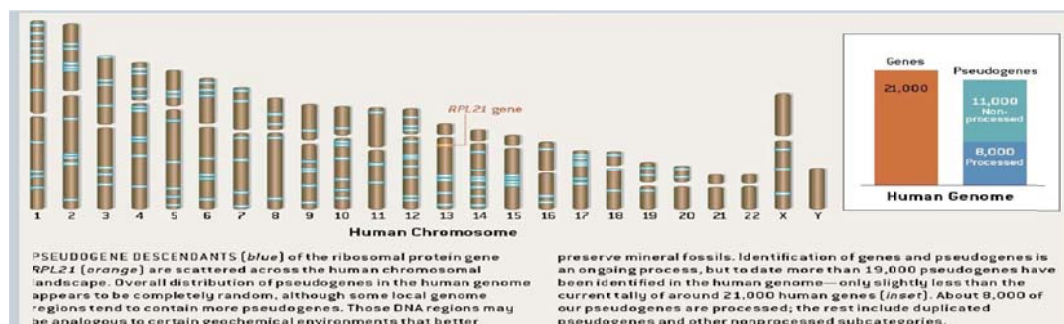
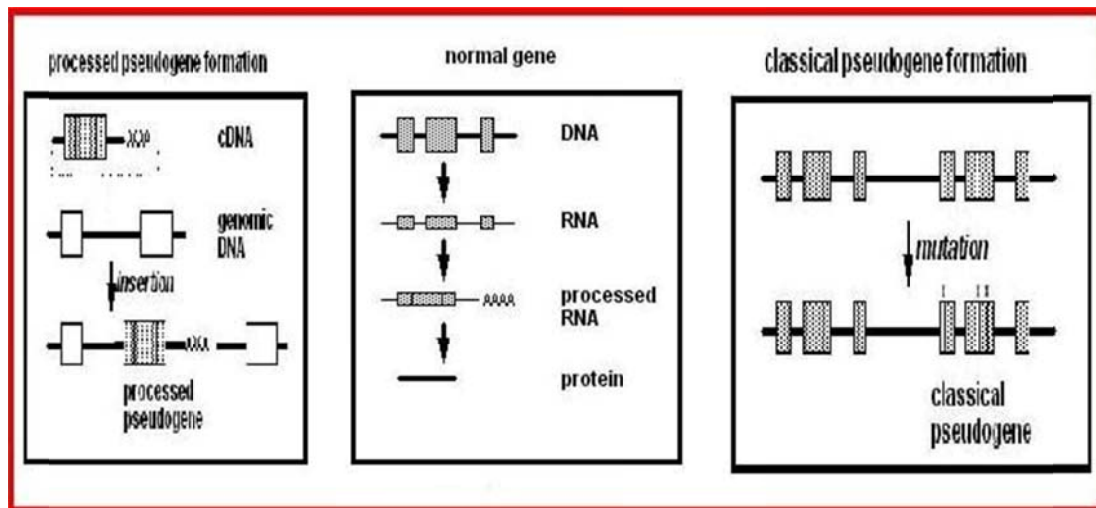


Figure 14-A: Pseudogenes in nuclear human genome



**Figure 14-B: Mechanisms of pseudogene formation**

Gerstein, M & Zheng, D. The real life of pseudogenes. *Sci Am* 295: 48-55 (2006).

### 7.3. Pyknons

In contrast to the solid concepts of classic genes functioning in traditional ways, the recent discovery of pyknons added new insights to the complexity of the functional organization of the human genome. Pyknons (**Figure 15**) are short non-coding DNA sequences about 20-22 nucleotides in length. They are widely distributed in the nuclear human genome in both the inter-genic and intronic regions of the genome, constituting about  $1/6^{\text{th}}$  of the human genome. This makes them the most frequent, variable-length DNA sequence motifs in the human genomes. Pyknons have a remarkable degree of structural conservation. Their presence in the 3' UTRs (untranslated regions) of genes may indicate a potential regulatory role in posttranscriptional processing and modifications of mRNA. Though they do not share in either protein synthesis or RNA transcription, pyknons are functional genetic elements associated with mediation of specific biologic cellular processes. They are putative factors implicated in susceptibility to some common human genetic disorders. Disturbed genomic regulation of function(s) of pyknons might underlie the development of this genetic susceptibility. The considerable size of pyknons in the genome coupled with the intimate functional relationship between them and many subtypes of microRNAs suggest a pivotal role played by pyknons, probably as global regulators of gene function. Some unique sequences of human genomes are designed from series of short template octamer sequences which are embedded into pyknon's sequences and represented by hundreds (up to thousands) of copies in a genome [**Gennadi V., 2009**]. The assumed regulatory roles of pyknons might be exerted via different mechanisms. The small nucleotide number of pyknons, similar to the small nucleotide number of most microRNAs species, elicits some questions regarding their origin in the genome, and the possibility that pyknons might represent **non-classic genes** or **transcriptional units** capable of directing and regulating synthesis of some microRNAs species for specific biological activities.

>3'UTR OF ENST00000245836|ENSG0000010196  
 -(162)-TTTCTTGATTTTTCAG-(104)-ATACTGATTTA  
 GTATT-(146)-TGGATGAAAAATATTT-(591)-CTCAC  
 CTGCTCACTGCA-(6)-GCCTTCTGGGTTCAAG-(2)-AT  
 TCCTGACCTCAAG-(19)-TCCCAAAGTGCTGGGATTA-  
 ATCCCAGCACTTTGGGA-(14)-GATCACGAGGTCAGG  
 ACT-(5)-GGCTGAGGCAGGAGAAT-GGTGTGAACCCG  
 AAAAT-(69)-ACTCCCATCCTAATAC-(137)-CGGTGG  
 CTGGGCGTGGTG-(8)-CTGTAATCCCAGCTACT-(5)-G  
 AAGAGCAAAACT-(58)-TTAGCCAGGCGTGGTG-(16)  
 CTCCAGTCTGGG-(2)-ACAGAGTGAGACCCCA-(50)-  
 TAACAACATTTTACAT-(37)-AGCAATTATTTTAAA-  
 G-(9)-TGTAGTCTGGCTACT-(15)-GGAGGATCGCTT  
 TACT-(23)-CAGTCTCACTGTGTTG-(4)-GGATGGAGT  
 GA-(15)-TTGGCCAGGCTAGTCT-(2)-AATTTCTGACC  
 GGTGAT-(8)-CTCGGCCTCACAAAGTG-CTGGGATTAC  
 T-(2)-AGACCAGCCTGGCCAA-(11)-CCTGTCTGTACA  
 GC-(1)-GAGGTTGCAGGGAGCC-(13)-CGCACTCCAG  
 GTGAAT-(10)-CAATAAAATACTATT-(10)GCACCTT

Figure 15: Pyknons

#### 7.4. Transposons

Transposons represent a unique feature of the genome of most living creatures. They represent one type of mobile genetic elements (**MGE**) which are sequences of the genome that can move from their original locations to other sites within the genome or make a copy of their sequence to be inserted in other parts of the genome. Other mobile genetic elements include plasmids, group II catalytic introns or ribosymes and bacteriophages. The movement of transposons from their site to another site happens via one of two mechanisms: the **replicative** and the **conservative** mechanisms. In the replicative mechanism, the transposon element replicates making a copy of itself and the new copy gets inserted in a new site of the genome thus leading to duplication of the transposon sequence. This category of transposons is named **Class I Retrotransposons**. The second group of transposons, **Class II DNA Transposons**, follows the conservative mechanism where the transposon detaches and moves, or transposes itself, from its original location on a specific chromosome to a new site on the same or on a different chromosome (**Figure 16**). In either case, insertion of a new segment into a normal segment leads to disruption of the integrity of the normal sequence. This phenomenon, referred to as **insertional mutagenesis**, is shared with many retroviruses. The resultant effect of this process depends on the site of insertion of the mobile element. If it occurs within nonfunctional inter-genic portions of the genome no harm is to be expected. However, insertional mutagenesis within functional genes results in variable degrees of damage depending on many factors, like the size and the site of the new insertion. It usually results in interruption of structural integrity of the gene and constitutes one major factor that underlies the occurrence of **spontaneous mutations** of the genetic material and pathogenesis of genetic disorders [Slotkin and Martienssen, 2007].

On the other hand, creation of new genetic combinations between receptor segments of the genome and inserted transposons might be considered, within the context of **evolutionary genetics**, as one mechanism for **genomic diversity** and **phenotypic evolution** if these new genetic combinations result in construction of new functional



genetic elements and begin their own expression. Formation of new metabolic networks and establishment of novel organized regulatory pathways are two possible mechanisms that can underlie the acquisition of new genetic phenotypes due to transposon activity.

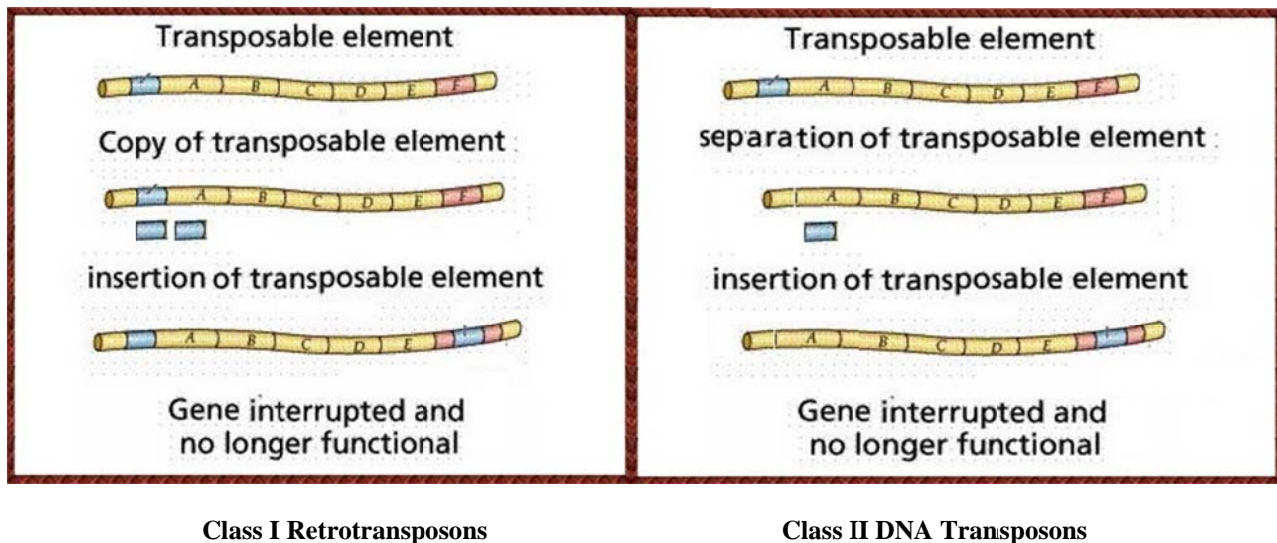


Figure 16: Transposons' movement and mechanisms of insertional mutagenesis

## 7.5. Telomeres

Telomeres (**Figure 17**) are specific noncoding, repetitive nucleotide sequences consisting of as many as 2000 repeats of the sequence (5' TTAGGG 3') located at the ends of linear chromosomes of most eukaryotic organisms. They protect the chromosome ends from being fused to each other and from fraying upon exposure to damaging agents. Over time, however, each cell division cycle results in loss of a portion of the telomere sequence leading to progressive shortening of the telomere because DNA replication cannot continue their duplication all the way to the end of chromosomes. If cells divide without telomeres, they would lose the ends of their chromosomes, and the important genetic information they contain. In human blood cells, the length of telomeres ranges from 8,000 base pairs at birth to 3,000 base pairs as people age and as low as 1,500 in elderly people. Each time a cell divides, an average person loses 30 to 200 base pairs from the ends of that cell's telomeres [Dan Eisenberg, 2011].

Consumption of telomere portions during cell division is partly corrected by resynthesis by a specific enzyme named telomerase reverse transcriptase. This enzyme is found only in certain types of cells which comprise germ line cells, embryonic stem cells and adult stem cells including cancer stem cells and their progenitor cells. The activity of the enzyme in these cell types explains many aspects of their biologic potentials. Prolonged and persistent synthesis of telomerase reverse transcriptase enzyme is a constant feature of most cancer cells and represents an important malignant phenotype of these cells underlying their ability to grow and divide indefinitely.

However, at a certain stage of somatic cell life cycle, no more telomere sequences could be lost and gradual deterioration of chromosome integrity ensues, leading ultimately to replicative senescence, enhanced aging and cell death. The role played by telomeres in regulating the number of cell divisions during the life span of the cell, referred to as the **Hayflick limit**, reveals the critical role played by telomeres in keeping genomic integrity and stability within safe functional limits all through the life span of the cell.

Though telomeres are found exclusively at the ends of linear chromosomes, **interstitial telomeric sequences** (ITSs) with their specific repeats of (5' TTAGGG 3') are found scattered throughout the human genome, particularly within the middle of chromosome 2 which contain pre-telomeric sequence, a telomeric sequence, an inverted telomeric sequence and an inverted pre-telomeric sequence. ITSs are often functionally important to the genome. The chromatin organization of telomeres can silence genes and has been linked to epigenetic modes of inheritance. Furthermore, difference classes of transcripts are derived from telomeres and their flanking repetitive DNA regions. These are involved in numerous cellular and developmental functions. It seems more likely that ITSs are sites where TTAGGG repeats have simply been added to chromosomes by telomerase enzyme and that many of these ITS sites are associated with distinct sets of proteins which have been linked to important functional roles within the genome, such as recombination hotspots.

Accumulating observations indicate that telomeres have important potential roles in many critical cellular processes. These processes include control of cell division, regulation of cell longevity, apoptosis, maintenance of optimal performance of stem cells and progenitor cells during early development and ensuring proper genomic replication of germ line cells during gametogenesis, among many others. The finding of a significant association between over-expression of telomerase enzyme and development of human cancers suggests new approaches to cancer therapy via combating this increased telomerase activity.

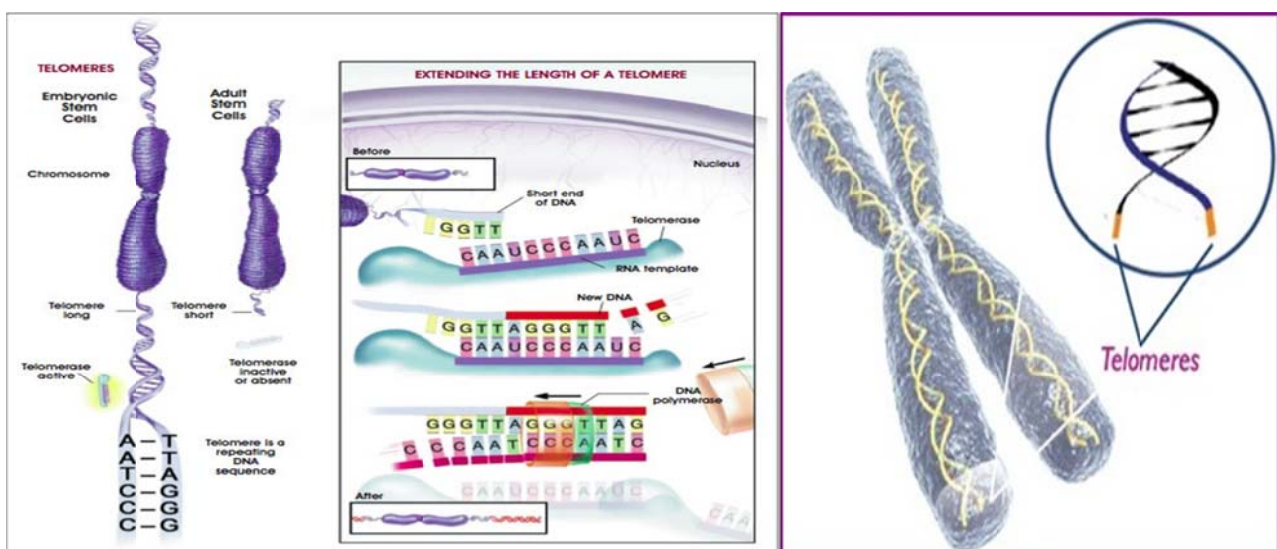


Figure 17: Telomeres

## 8. Structure of human genes

The gene, which is the functional unit of the genome, is a unique linear segment composed of a specific number of nucleotides arranged in a characteristic specific sequence, along one strand of DNA, the coding or active strand. Each gene has a specific site on the DNA constituting a particular chromosome called the **gene locus** which is characteristic of each gene. So, a gene might occupy a specific locus on the short arm or the long arm of the chromosome (**Figure 18**).

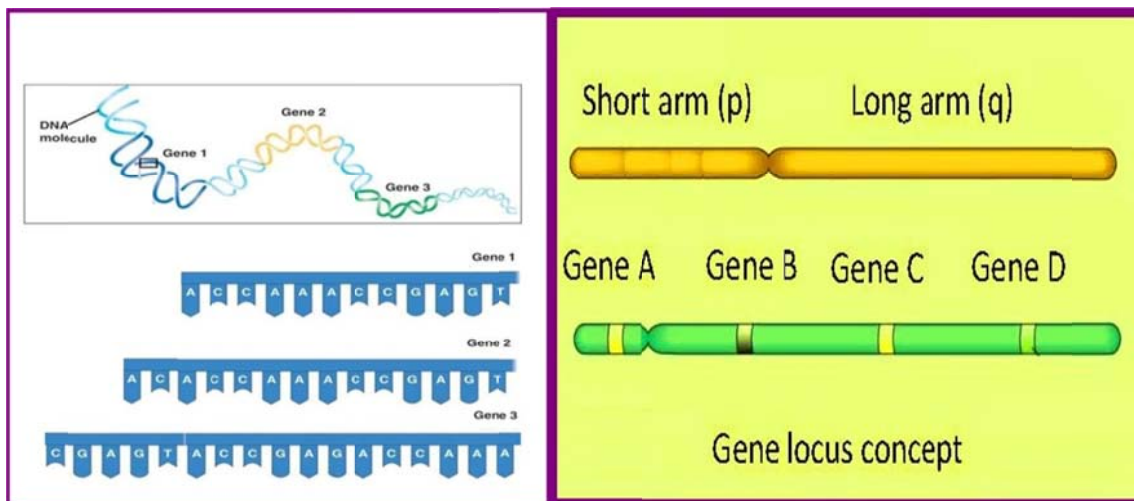


Figure 18: Linear arrangement of genes and the gene locus concept

All genes have the same basic structure, being composed of a long linear piece of DNA composed of the 4 nucleotides (A,G,C,T), but in varying numbers and a peculiar arrangement characteristic of each gene. Some genes are small sized genes formed only of few hundred nucleotides, e.g. globin genes, while others consist of much larger numbers, reaching up to 2.4 million nucleotides in case of the dystrophin gene. The specific arrangement of the nucleotides of the gene imparts to each gene its functional specificity. Functionally, structural genes differ from each other by the structure and nature of the protein(s) synthesized under their control, which is determined by the specific arrangement of the nucleotides of the gene.

## 9. Functions of human genes

As referred to previously, human genes comprise three major functional categories: **structural genes** which are directly involved in, and responsible for, protein synthesis, **regulatory genes** which regulate the functions of structural genes and **master genes** which monitor and control vital cellular activities. Structural genes constitute the major proportion of genes in the genome. They function in strict harmony to regulate the synthesis of proteins, both structural proteins and catalytic proteins or enzymes, necessary for mediating all metabolic and life processes in the cell. For simplification and with very few exceptions, the structural gene is functionally divided into three main parts: the **promotor** of the gene, or the part of the gene responsible for switching on the gene to begin working and switching off the gene to stop working, the end or **terminal portion** of the gene and the main body or

the **transcribing portion** of the gene lying in between the gene promotor and the end of the gene and responsible for the actual work of protein synthesis. All parts of the gene are composed of the same four nucleotides but in varying numbers and in peculiar sequence characteristic of each gene. The transcribing portion of the gene is functionally recognized as being divided into alternating segments, the **exons** or the base sequences that participate in defining the protein coded by the gene and the **introns** or the base sequences that, with few exceptions, do not participate in defining the gene product. Each gene has its specific structural characteristics, or internal anatomy, that distinguishes it as regards the number of exons and introns of its transcribing portion, the length of the terminal portion of the gene and the base sequence of the promoter area (**Figure 19**).



**Figure 19: General scheme of functional organization of genes**

## **10. Stages of gene function**

1. Gene activation
2. Transcription (synthesis of mRNA)
3. Post-transcription modifications of m-RNA
4. Translation (synthesis of protein)
5. Post-translation modifications of proteins.

### **10.1. Gene activation**

Protein synthesis in the cell occurs under strict control of structural genes and is a multistage process. It involves an intricate cascade of synchronized and interrelated processes beginning with activation of, or switching on, the gene to begin functioning. This activation happens through binding of specific activation, or transcription factors, mostly nucleoproteins or miRNA molecules, to specific sequences of the promotor of the gene. This activation process is regulated by regulatory genes that sense the need for the gene product according to the metabolic demands of the cell.

As the main dogma of molecular biology entails, the **genome** controls the synthesis of the **proteome** that mediates all cellular activities. An intermediate step in this process as regards function of structural genes is synthesis of an intermediate mRNA with the same information contained in the gene and necessary for protein synthesis. This information is referred to as the **genetic code** which is the specific



bioinformatics database embodied in the gene and represented by the peculiar arrangement of its bases. The genetic code is designed so that each **three bases in sequence**, known as the **codon**, along the transcribing portions of the gene, the exons, define a **specific amino acid** in the synthesized polypeptide chain. The codon might be considered as the functional unit of the gene since each codon within the exon defines an amino acid in the protein. Sixty four (64) codons exist in nature, sixty one (61) of them define amino acids while the remaining three (3) codons, stop or termination codons, do not define any amino acids and are used to locate the end point of protein synthesis, after the proper protein with the exact number of amino acids has been synthesised. Some amino acids are defined by more than one codon, e.g. six different codons can define leucine. This phenomenon is referred to as **degeneracy of the genetic code** and is an important **anti-mutation mechanism** since it allows occurrence of point mutations in the codon of the gene without change of the amino acid in the synthesized protein (**Figure 20**).

The Genetic Code							
TTT	Phe	TCT	Ser	TAT	Tyr	TGT	Cys
TTC	Phe	TCC	Ser	TAC	Tyr	TGC	Cys
TTA	Leu	TCA	Ser	TAA	STOP	TGA	STOP
TTG	Leu	TCG	Ser	TAG	STOP	TGG	Trp
CTT	Leu	CCT	Pro	CAT	His	CGT	Arg
CTC	Leu	CCC	Pro	CAC	His	CGC	Arg
CTA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
CTG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
ATT	Ile	ACT	Thr	AAT	Asn	AGT	Ser
ATC	Ile	ACC	Thr	AAC	Asn	AGC	Ser
ATA	Ile	ACA	Thr	AAA	Lys	AGA	Arg
ATG	Met*	ACG	Thr	AAG	Lys	AGG	Arg
GTT	Val	GCT	Ala	GAT	Asp	GGT	Gly
GTC	Val	GCC	Ala	GAC	Asp	GGC	Gly
GTA	Val	GCA	Ala	GAA	Glu	GGA	Gly
GTG	Val	GCG	Ala	GAG	Glu	GGG	Gly

1. The genetic code is the information contained within the specific arrangement of the bases of the structural gene that determines the amino acid sequence of the protein coded by the gene.

2. It is designed so that each three bases in sequence (Codon) along a specific position of the gene define a specific amino acid at a specific position along the protein produced by the gene.

3. There are 64 known codons: 61 functional codons and 3 termination or stop codons that determine the end of the synthesized protein.

4. The genetic code is universal, it is used by all known forms of life. However, minor variations do exist for genetic codes of mitochondria, bacteria, mycoplasma candida and archaea.

Figure 20: The genetic code

## 10.2. Transcription

Synthesis of mRNA based on DNA base sequence, or transcription, is a main step in gene function since the genes reside in the nucleus while protein synthesis occurs in the cytoplasm, and an intermediate carrier that can carry the genetic information of the gene in the nucleus to be used for protein synthesis in the cytoplasm is needed. Nuclear DNA, structurally organized as chromosomes, is a large and heavy structure

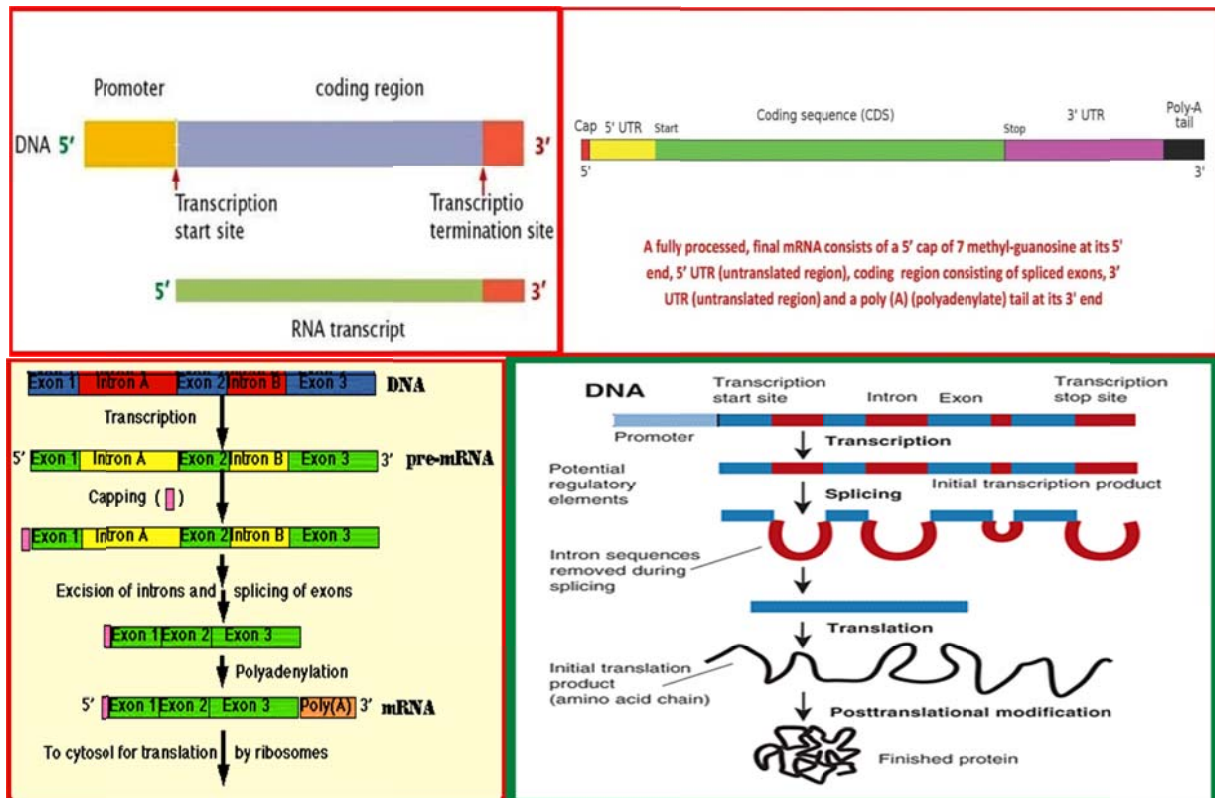
that can't either pass through the nuclear membrane or exists in the cytoplasm because it might be damaged by enzymes and reactive metabolic radicals. The primary mRNA synthesized in the nucleus consists of all the bases of all exons and introns of the gene, but in a complementary sequence, i.e. Cytosine in place of Guanine, Guanine in place of Cytosine, Adenine in place of Thymine and Uracil, not Thymine, in place of Adenine. Primary mRNA undergoes many **post-transcription modifications** before it turns to final mRNA.

### 10.3. Post-transcription modifications of mRNA

These structural modifications of primary mRNA transcript are necessary for maturation of the mRNA to render it a final transcript ready for translation. They include the following steps:

1. Excision and removal of introns and joining or splicing of exons so that the final, or mature, mRNA consists only of sequences complementary to those of the exons, since introns do not participate in protein synthesis. Retaining one or more introns in the mature mRNA by selective alternative intron excision allows a single gene to make different mRNA transcripts, each one capable of being translated to a different protein. This process of **selective alternative excision of introns** clarifies the ability of some genes, notably immunoglobulin genes, to produce large numbers of proteins by the same gene and explains the marked discrepancy between the estimated number of human genes, about 20000 – 40000 genes, and the number of proteins produced by them, estimated to lie between 400000 and 4 million proteins.
2. Addition of a 5' 7-methylguanylate cap to the first base of the mRNA to protect it from degradation by RNA exonucleases in the cytoplasm. The 5' **capping** probably has other functions like regulation of exit of the mRNA from the nucleus, attachment of the molecule to specific factors in the ribosome and promotion of proper translation of the molecule by the ribosome. The cap also plays an important role in messenger RNA degradation which is a critical process for regulating the turnover of RNA in the cell. RNA degradation is initiated by **deadenylation** of the polyadenylated tail followed by **decapping** of the 7-methylguanylate cap through hydrolysis, thus exposing the 5' end to 5' to 3' exoribonuclease activities.
3. Addition of a long stretch, 100-300, of adenosine monophosphate molecules, poly (A) tail, at the 3' end of the mRNA. This process, known as **polyadenylation**, is necessary for stabilizing the mRNA molecule and protecting it from enzymatic degradation in the cytoplasm. Also, polyadenylation is needed for transcription termination, export from the nucleus and regulation of translation. **Selective alternative polyadenylation** by addition of the polyadenylate tail at different sites of the mRNA molecule, similar to selective alternative intron excision, permits for the production of many different transcripts of mRNA varying in their lengths and each is capable of being translated to a different protein. This process of selective alternative polyadenylation also explains the marked ability of the genome to regulate

the synthesis of such a huge number of proteins that constitute the proteome (**Figure 21**).



**Figure 21: Post-transcription modifications of mRNA**

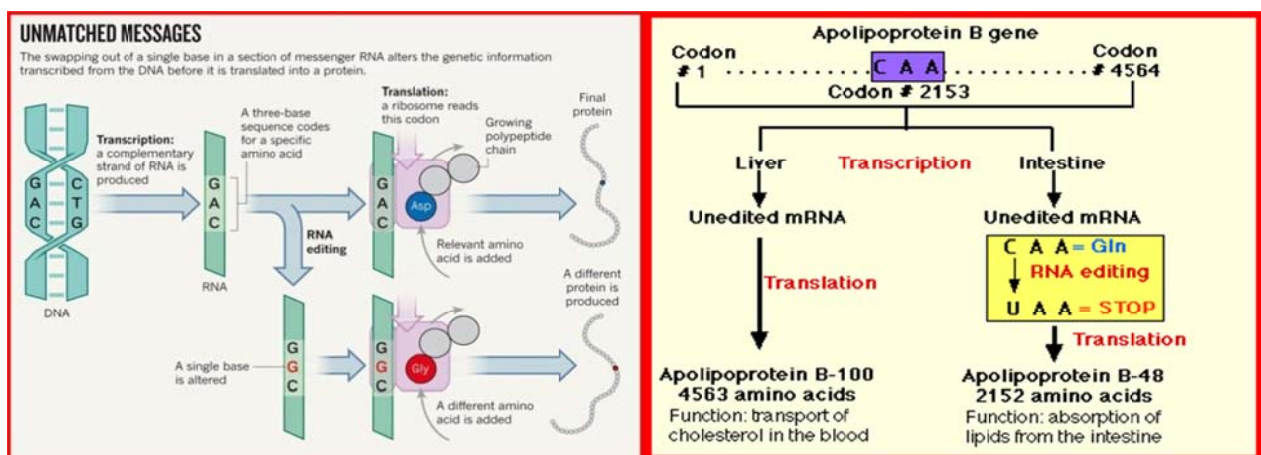
**4. Circularization** of mRNA entails formation of a circular mRNA molecule by binding the 5' cap region to the 3' polyadenylate tail by attachment to certain proteins that act as translation factors. Circularization is thought to promote cycling of ribosomes on the mRNA leading to efficient decoding of the molecule and synergistic enhancement of translation. It also acts as a proof reading system to ensure that only intact circularized mRNA molecules are translated by the ribosomes, thus excluding partially degraded mRNA that do not have 7-methylguanylatecap or poly-A tail, from being translated to defective proteins.

**5. Editing** of mRNA represents an important post-transcription modification of mRNA that allows the replacement of one base of the molecule by another one. Accordingly, a transcript of RNA with base sequence differing from the template gene is produced. This process is mediated by a sophisticated system, the **editosome**, including a peculiar species of miRNA called guide RNA (gRNA) which has a specific structural configuration and uses different editing mechanisms, including insertion, deletion or deamination of bases. There is accumulating evidence that RNA editing occurs extensively in many types of human cells.

The most obvious example of mRNA editing in human is the editing of the mRNA transcript of the Apolipoprotein B gene by turning a functional codon within the

transcript to a stop codon leading to synthesis of a full length protein in the liver (apolipoprotein B-100) consisting of 4536 amino acids and functioning as a transporter of cholesterol in blood, and synthesis of a different shorter truncated protein (apolipoprotein B-48) composed of 2152 amino acids that functions in intestinal absorption of lipids. Other examples of editing of human mRNA include different proteins acting as receptors for neurotransmitters in the brain and proteins that constitute structural units of cell membrane transport channels.

The exact significance of editing of mRNA is still a matter of debate. However, editing might be thought of as a genetic **repair mechanism** at the level of RNA, compared to DNA repair system, that can repair some point or small mutations of the mRNA transcript. The alteration of mRNA sequence due to editing leads to effects similar to those resulting from selective alternative excision of introns and selective alternative polyadenylation of mRNA, since it allows the synthesis of a different protein by the edited molecule. RNA editing offers an additional explanation for the voluminous quantitative difference between the number of functional genes of the genome and the number of proteins that make up the proteome (**Figure 22**).



**Figure 22: Post-transcription editing of mRNA and guide RNA**

## 10.4. Translation

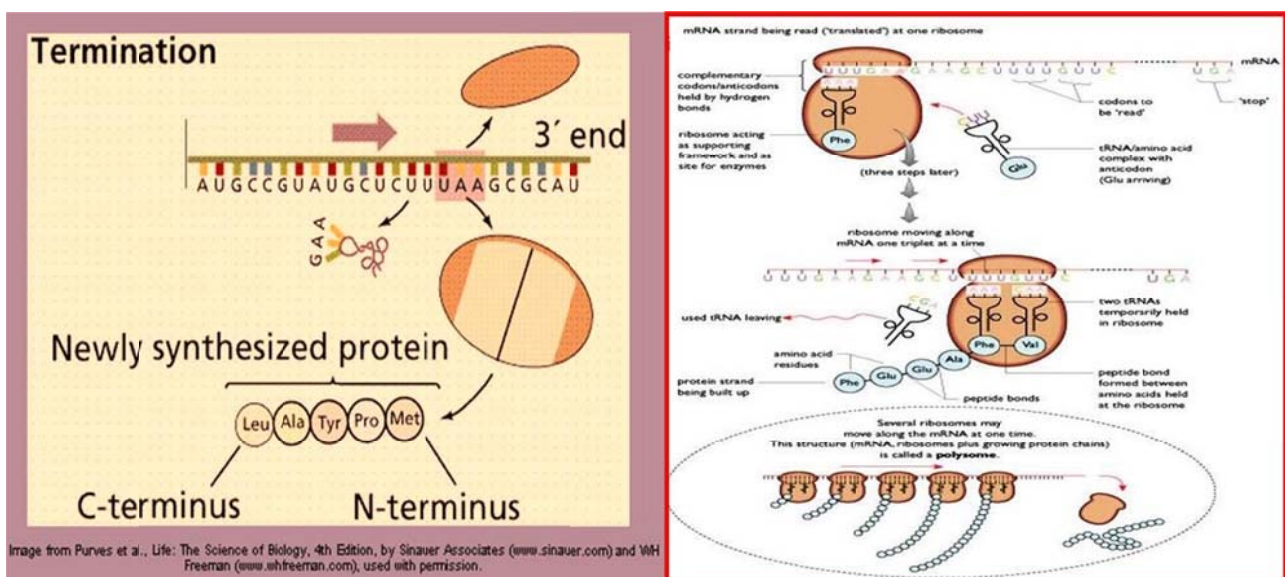
Translation is the actual process of protein synthesis based on the genetic information contained within the specific codon sequence of mRNA. It entails translation of gene language, in the form of codons, to protein language, in the form of amino acids. Translation is a very sophisticated process accomplished in the cytoplasm by the synergistic and cooperative action of **mRNA**, **ribosomes**, **rRNA** (ribosomal RNA), **tRNA** (transfer RNA) with the participation of tens, sometimes hundreds, of protein and non-protein **translation factors** that mediate the associated steps of the multiple sequential mechanisms of the translation process.

Translation happens by attachment of the ribosome with its rRNA to the mRNA. With the participation of the codon recognition segment of rRNA and the anticodon segment of tRNA, decoding, translating or recognizing the amino acid defined by the first codon of mRNA results in bringing the specified amino acid from the cytosol of



the cell, by its specific carrier tRNA amino acid attachment site, to the site of protein synthesis in the cytoplasm. The ribosomal complex moves forward along the mRNA to repeat the same steps with the second codon of mRNA, and the second specified amino acid is brought next to the first amino acid where a peptide bond is formed between both amino acids.

With repetition of the decoding, recognition or translation process of codons of mRNA, sequential addition of the specified amino acids leads to progressive lengthening of the synthesized polypeptide chain till the ribosomal complex reaches the end of the coding region of mRNA where it finds a stop or termination codon denoting the end of the coding region of the transcript. Recognition of a stop codon by the ribosomal complex means that no more amino acids are to be added to the polypeptide chain and that the proper protein with the proper number of amino acids, as dictated by the original genetic information of the gene, has been synthesized. This recognition signal marks the end of the translation process where separation of the ribosome from mRNA occurs followed by consequent degradation of the mRNA with metabolic recycling of its nucleotide content within the metabolic pool of the cell. Repetitive cycles of translation of the same mRNA continue till sufficient amounts of the protein required for metabolic activities of the cell are produced (**Figure 23**).



**Figure 23: Protein translation**

## 10.5. Post-translation modifications

Post-translation modifications refer to changes of the structural configuration of newly synthesized proteins. The majority of newly synthesized proteins have to undergo specific structural modifications, e.g. folding, to become functionally active in accord with the concept of **form-function relationship** that governs functional efficiency of protein biomolecules. These conformational changes or modifications of protein structure are critical for most proteins to confer upon them physiological potency since the vast majority of newly formed proteins, with rare exceptions, are

unable to perform their physiological functions unless they undergo these modifications. Failure of completing these structural modifications leads to production of defective proteins unable to mediate their roles in cellular activities, and underlies the development of a large number of genetic diseases.

New proteins are synthesized in the cytoplasm as long polypeptide chains. This basic structural configuration is referred to as **primary protein structure**. Multiple folding of the polypeptide chain along its length leading to formation of a new functional shape, e.g.  $\alpha$ -helices,  $\beta$ -pleated sheets, tight turns and  $\beta$ -bulges, due to hydrogen bonding between specific amino acid sequences of the chain, is one mechanism used to activate proteins and demonstrates the mechanism of acquisition of a **secondary protein structure**. Some proteins need to acquire a three dimensional configuration, by intra-molecular binding and interactions of some of its amino acid sequences, in order to form distinct functional domains of the protein to become physiologically active component. This complex configuration represents the **tertiary protein structure**. Still, some proteins, e.g. enzymes, must have a more complex configuration to be physiologically active. **Quaternary protein structure** is attained via a wide spectrum of modifications of the protein that include addition of another protein, addition of a non-protein moiety like vitamins, heavy metals or reactive chemical groups that act as cofactors or prosthetic groups for most proteins that act as enzymes (**Figure 24**).

Post-translation modifications of proteins span, nearly, an endless spectrum of changes of the protein structure. An important aspect of this spectrum includes chemical alteration of proteins by binding or reacting with biochemically and biologically active chemical groups, e.g. phosphorylation, glycosylation, acetylation, methylation and sulfation. The spectrum of post-translation modifications includes also addition of water molecules that become closely bound around proteins. These molecules play critical roles in stabilizing the protein folding and they greatly influence substrate binding and catalytic potential of enzymes. This important role of water in maintaining structure and enhancing function of proteins explains the drastic consequences that result from disturbances of water metabolism in cells, for instance complications following states of dehydration or over hydration.

In view of the necessity of post-translation modifications of newly synthesized proteins for proper functioning, it seems quite plausible to consider this process as an integral stage of gene function, because they are indispensable for the completion of the preceding stages of gene function. Additionally, these modifications, with very few exceptions including spontaneous inherent folding of some proteins, are conducted by large numbers of enzymes that mediate these modifications under strict regulatory control of many genes responsible for these modifications.

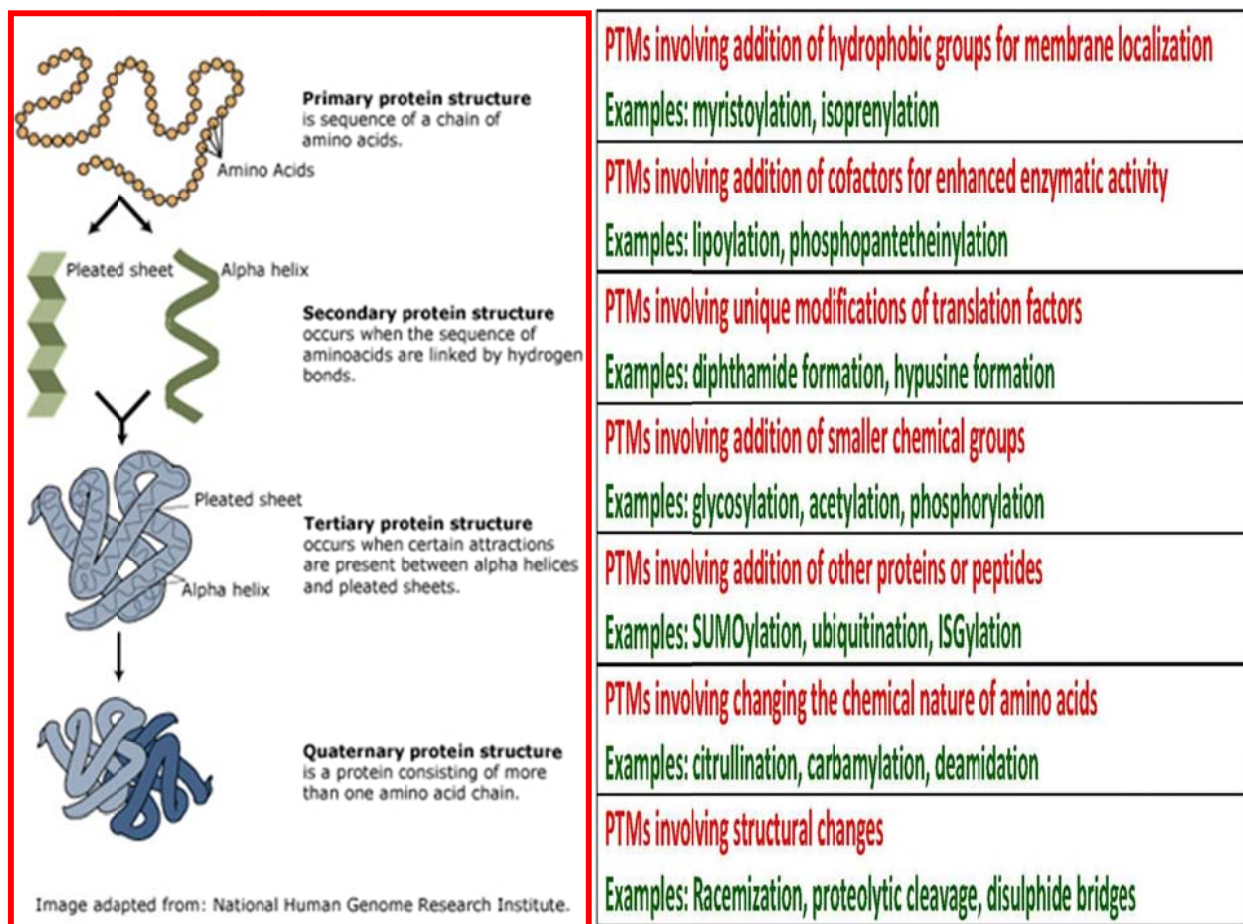


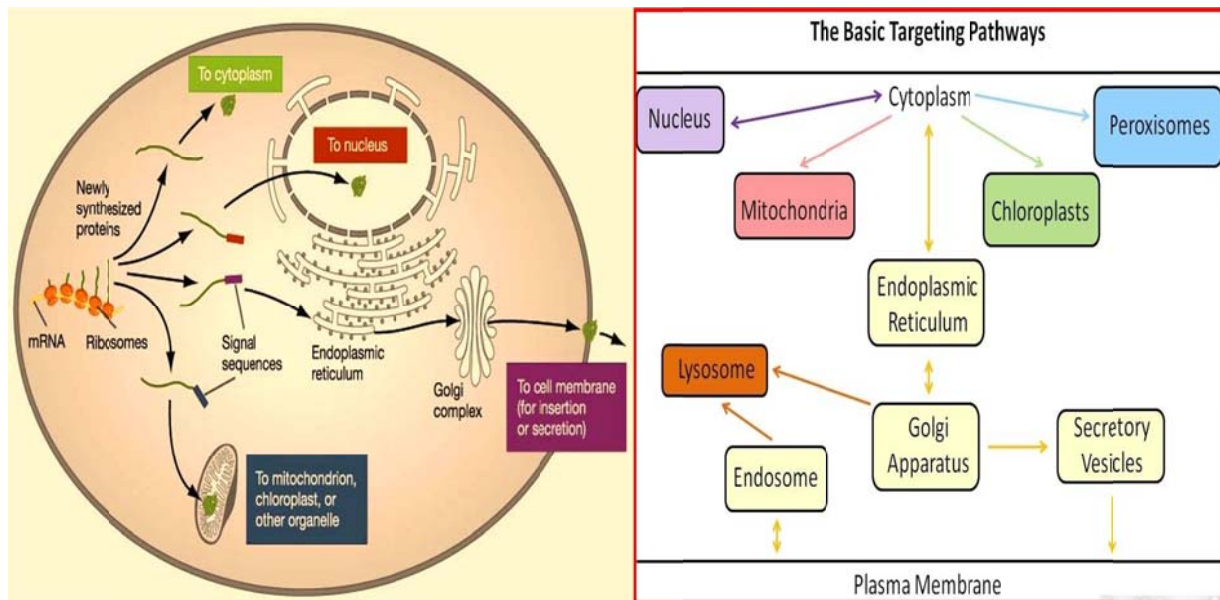
Figure 24: Post-translation modifications of proteins (PTMs)

## 11. Post-translation trafficking of proteins

Life activities within living cells are mediated by proteins. Two major classes of proteins can be recognized within this functional context: **structural proteins**, like cytoskeleton proteins, and **catalytic proteins** or **enzymes** that conduct and regulate all metabolic networks encompassing biological processes within the cell. Proteins are highly specialized biomolecules. Their functional specialization is intimately dependent on their proper localization at their targeted sites of action inside the cell. Accordingly, newly formed proteins within the cytoplasm have to be transported from their site of synthesis and directed to their sites of biological action, e.g. insertion in cellular membranes or catalysis of metabolic activities inside the mitochondria.

Post-translation trafficking of proteins (**Figure 25**) refers to the dynamic processes that follow synthesis of new proteins, aiming at their proper localization within the cell compartments. Trafficking is critical for proper functioning of proteins. Precise targeting of proteins depend on synthesis of specific factors, mostly short amino acid sequences or chemical molecules like mannose-6-phosphate, that direct the transport of the protein to its exact destination. Post-translation trafficking of proteins involves active participation of the endoplasmic reticulum and the Golgi apparatus. Passage through one or both of these organelles is necessary for many proteins to become active biomolecules or to get ready for attachment to their specific recognition signal

molecules needed for trafficking to their proper sites. The rough endoplasmic reticulum (RER) is an integral part of the protein targeting pathway. Proteins that pass through it and exit from there are marked with signal sequences that work as address label directing the proteins to their destination. In the absence of protein targeting signals, newly formed proteins remain functionless at their sites of synthesis in the cytoplasm.



**Figure 25: Post-translation trafficking of proteins**

Defects in trafficking processes can result in disturbed localization of the protein to its target site with resultant functional deficiency of its biological and/or metabolic activities. Many genetic diseases are caused by failure of targeting properly synthesized proteins from their sites of synthesis to their sites of action.



## Part 2: Basic concepts of medical pathgenetics

### 12. Pathogenetics

The term pathogenetics has been coined to refer to the multiple and different mechanisms involved in pathogenesis of genetic diseases. It comprises the study of **mutagens**, the study of **mutations**, the study of the different **pathogenetic mechanisms** that result from disordered gene functions secondary to change of gene structure, the study of the **pathophysiological alterations** in cellular functions due to disturbances of the metabolic-regulatory networks that mediate and control these functions and, finally, the study of pathogenesis of genetic diseases.

#### 12.1. Mutagens

Mutagens are factors that can cause mutation of the genetic material. Mutagens are plentiful in our life. In fact, we live in a mutagenic world, and with the exception of pure water any compound in our environment is either an obvious mutagen or can be a potential mutagenic factor under favorable circumstances.

#### 12.2. Classification of mutagens

Mutagens can be classified according to their nature or according to their recognizable pathogenetic and pathological effects on exposed cells (**Table 7**).

A. According to their **nature**, mutagens are classified into four main categories:

1. **Chemical mutagens**: these compounds are innumerable in the environment and include, for instance, thousands of organic compounds, asbestos, insecticides, herbicides, heavy metals, etc.
2. **Physical mutagens**: these include particulate radiations like X ray, alpha particles, UV waves at 2800 Å wavelength, solar radiation, thermal effects and mechanical agitation of nucleic acids.
3. **Biological mutagens**: these include living microorganisms like some viruses (cytomegalovirus, rubella virus and herpes virus) and toxoplasma gondii.
4. **Changes of energy states** of hydrogen bonds between complementary bases of DNA.

B. According to their **effects**, mutagens are classified into four main categories:

1. **Non-specific mutagens**.
2. **Carcinogens** are mutagens that induce malignant transformations in affected cells.
3. **Clastogens** are mutagenic agents that can induce chromosome breaks in affected cells.
4. **Teratogens** are mutagenic factors that cause congenital malformations in exposed fetuses.

**Table 7: Classification of mutagens**

Classification of mutagens			
According to effect	According to nature	Examples	Mechanisms and effects
<b>1. Carcinogens</b> Mutagens that induce malignant transformation and development of cancer  <b>2. Clastogens</b> Mutagens that cause breakage of chromosomes  <b>3. Teratogens</b> Mutagens that cause congenital malformations in developing embryo  <b>4. Non-specific mutagens</b> Mutagens that cause non-specific mutagenic damage to the genetic material.	<b>1. Physical mutagens</b>	1. Particulate radiations: X- ray, alpha particles, neutrons, positrons, gamma rays, protons. 2. UV waves 3. Solar radiation & electromagnetic waves 4. Thermal exposure 5. Mechanical agitation	1. Breakage of nucleic acids 2. Chromosome breaks 3. Oxidation of deoxyribose 4. Formation of peroxides 5. Deamination of bases 6. Dehydroxylation of bases 7. Ionization and formation of ion tracks/ion clusters 8. Generation of free radicals 9. Provocation of carcinogenesis 10. Formation of pyrimidine dimers 11. Denaturation of DNA
	<b>2. Chemical mutagens</b>	Organic hydrocarbons, acridine derivatives, manganese, nitrous acid, hydroxylamine, base analogs, sulfonate compounds, insecticides, asbestos, herbicides, heavy metals.	1. Base transition 2. Base transversion 3. Deamination of bases 4. Generation of free radicals 5. Interruption of base Sequence by intercalation 6. Provocation of carcinogenesis 7. Cytopathic effects following mitDNA damage 8. Alkylation 9. Adduct formation
	<b>3. Biological mutagens</b>	1. Viruses: rubella virus, cytomegalovirus, HIV and herpes virus. 2. Toxoplasma gondii.	1. Insertional mutagenesis 2. Provocation of carcinogenesis
	<b>4. Changes of energy states of bonds between bases</b>	Tautomerism, hydration of bases	Base transition
	<b>5. Metabolic derangements</b>	1. High acidity 2. Dehydration 3. Peroxidation	Depurination of nucleic acids Provocation of aberrant DNA replication.

The pathogenetic effects of mutagens and the resulting pathological alterations follow disturbed functions of mutated genes. Many types of mutagens have the ability to induce all types of mutagenic damage to the cell. Particulate and non-particulate radiation, for example, are effective carcinogens capable of causing malignant transformation of affected cells. They can also cause chromosomal gaps and breaks. In addition, teratogenic effects of developing fetuses following teratogenic exposure of pregnant women to radiation is a well known hazard of overexposure to radiation during pregnancy. This global mutagenic effect is also shared by many other types of mutagens like mutagenic viruses and chemicals that can induce malignant transformation, chromosomal breaks and teratogenic malformations in exposed subjects.

### 12.3. Factors affecting actions of mutagens

The pathological effects of mutagens are determined by many different factors. Each of these factors has its relative mutagenic potential depending on the circumstances of its action. Congenital malformations caused by teratogens, for instance, depend largely on **timing of exposure** of the developing embryo or growing fetus to their effects. Exposure to teratogens before or after periods of embryonic or fetal growth sensitive to their specific teratogenic effects might result in no or minimal harm. The **dose of exposure**, whether of a chemical or radiological nature, is also an important factor in determining the resultant mutagenic effects. This dose-threshold effect might act in different synergistic ways. A larger exposure dose will quantitatively cause more damage to wider portions of the genetic material, thus resulting in widespread mutations of the genome. It might also cause damage to genes responsible for repair of mutated DNA, thus preventing repair of mutated genes. In addition, direct non-genetic destructive effects of cellular compartments, of blood supply to cells or of extracellular environment of affected cells will both hasten and increase the resulting damage.

The **genetic constitution**, or genetic background, of exposed subjects plays a critical role in determining the sensitivity to mutagenic factors as well as the extent of damage following exposure to their effects. The outstanding examples in this respect are the chromosome breakage syndromes, sometimes referred to as DNA or genetic instability syndromes. These diseases include xeroderma pigmentosum, ataxia telangiectasia, Nijmegen syndrome, Bloom syndrome, Fanconi anemia and some other similar diseases. Subjects affected with these genetic disorders have deficient and/or defective **DNA repair systems** necessary for repairing mutations arising during DNA replication or following exposure to mutagens. Exposure to mutagenic factors in doses that have no effects in normal subjects causes drastic genetic alterations in these patients due to their defective genetic constitution.

### 12.4. Mutation

Mutation signifies any uncoded or unprogrammed permanent structural alteration of the genetic material at any of its organizational levels. These levels comprise a wide spectrum beginning with a whole nucleotide or just part of it (base, sugar, phosphate),

DNA, RNA, genes, chromosomes, mitochondrial DNA (mtDNA), up to the whole genome.

The effects of mutation differ widely according to many factors. These factors include the nature and target of the mutagenic factor causing the mutation, the timing and magnitude of the resulting damage, the genotype or the genetic constitution of the affected individual and the balance between synergistic mutagenic effects and **anti-mutation mechanisms** of the genetic material. The damaging effects of mutation are attributed to the defects they cause in functioning of the mutated genes. Since synthesis of proper gene products, necessary for mediating cellular activities, depends primarily on integrity of the genetic information embodied within the specific base sequence of the gene, changes or mutation of the exact number or the peculiar arrangement of these bases is expected to result in disturbed gene function. This disturbance might express itself as production of structurally defective gene product, deficient synthesis of enough product or disturbance in regulatory mechanisms responsible for monitoring, harmonizing and controlling gene functions.

### 12.5. Types of mutation

Mutation can be classified to many different types according to the parameter of classification (**Table 8**).

**Table 8: Classification of mutation**

Classification of mutation
1. Induced versus spontaneous mutation
2. Nuclear versus mitochondrial mutation
3. Somatic versus germinal mutation
4. Static versus dynamic mutation
5. Pathological versus non-pathological mutation
6. Persistent versus reversible mutation
7. Point, Small, Gross, Genomic mutation
8. Base, Sugar, Phosphate group mutation

#### 12.5.1. Induced versus spontaneous mutation

Induced mutations are structural alterations of the genetic material that follow exposure to known mutagens. Spontaneous mutations, on the other hand, are mutations that occur without recognizable causes. However, at least two possible potential causes of spontaneous mutation of the genome can be delineated. **Transposon activity** represents a major potential cause of spontaneous mutations. The movement of transposons results in insertional mutagenesis and disruption of targeted genes or other functional genetic elements. Disturbed gene function with resultant deficiency of gene product, production of defective protein/RNA or loss of genetic regulatory mechanisms exerted by mutated genes follows transposon-induced mutations.

**Changes of energy states** of hydrogen bonds between complementary bases of DNA represent another potential cause of spontaneous mutation. Alterations resulting from changes in **electron resonance** and/or proton transfer or relocation between the bonds can lead to tautomeric shift or tautomerization and momentary change of one base to another base, thus inducing point mutation, transition mutation. The actual mechanisms underlying the occurrence of these changes are still largely undefined, precisely. Analysis of DNA base dynamics according to quantum, rather than to Newtonian, principles seems essential for clarifying these mechanisms, and the concept of **quantum biology**, or studying the structure and function of the genetic material based on principles and rules of quantum mechanics, would certainly prevail due to the progressive increase in our knowledge of the structure and function(s) of our genome and our proteome.

**Replicative errors** occurring during cell division and replication of the DNA are, arbitrarily, considered as one source of spontaneous mutations. These errors occur, and persist, due to many causes including failure of the **proof reading system** of DNA replication or failure of the **DNA repair system** to repair or correct the errors. It is estimated that during every somatic cell division cycle of nearly 20 hours duration, about 12000 point mutations occur along the whole genome. The vast majority of these replicative errors are repaired by the DNA repair system of the cell. If some of these errors escape detection and repair they will persist and be transferred to the genome of daughter cells and become **fixed mutations**.

### 12.5.2. Nuclear versus mitochondrial mutation

Nuclear mutations are mutations that affect the nuclear genome, and mitochondrial mutations are mutations that affect the mitochondrial genome. Mutations of nuclear genes account for occurrence of most genetic diseases. Due to the marked differences between the specific structural organization of each genome, each type of these mutations has its own characteristic features. **Nuclear mutations** may affect one or both alleles of a structural gene on an autosome or on a sex chromosome, thus giving rise to any of the traditional genetic disorders: autosomal and sex linked dominant and recessive diseases. The presence of most nuclear genes as pairs of alleles derived equally from both parents, paternal and maternal alleles, allows for this classification of genetic diseases due to nuclear gene mutations to be formulated clearly. **Mitochondrial mutations**, on the other hand, have different features due to the peculiar nature of the mtDNA. Each mtDNA molecule consists of 37 genes and exists inside the mitochondrion in large numbers, tens to thousands, depending on the metabolic activities of the cell and its demand for ATP. Thus, the copy number of each mtDNA molecule, and of mitochondrial genes is extremely huge in metabolically active organs, e.g. brain, retina, heart, liver, kidney and skeletal muscles. Mutations of mitochondrial genome can affect a considerable portion of its genes without affecting its efficiency in mediating its functions. It is estimated that mitochondrial dysfunction begins to occur when approximately 80 % of its genome

gets mutated. This **threshold effect** does not apply for nuclear mutations where mutation of one single allele can cause a serious disease, as is the case for many autosomal dominant disorders.

Mitochondria in somatic cells are derived from maternal mitochondria in the ovum. The sperm, with rare exceptions, does not offer any mitochondrial genes to the zygote, because sperm mitochondria are present in the neck piece of the sperm, which exists next to its head pro-nucleus and which do not participate in fertilization. Accordingly, mitochondrial disorders result from mutations of maternal mitochondria and are transmitted by carrier or affected mothers to all their offspring. This difference underlies the specific features that characterize the nature and the inheritance pattern of mitochondrial diseases.

### 12.5.3. Somatic versus germinal mutation

Somatic mutations refer to mutations of the genome of somatic cells. Since somatic cells do not participate in fertilization or determination of the genetic constitution of the offspring they can't be transmitted to the offspring, i.e. they are not heritable mutations. They can be transmitted only to the daughter cells of mutated cells upon their division. The **effects of somatic mutations** depend on many factors including the type of cell, the genetic constitution of affected cells, selective targeting of nuclear and/or mitochondrial genome and the mutation burden of the cell. According to the interactive processes involving these factors, somatic mutations might result in **cell death** if the mutation-induced pathophysiological alterations of the cell exceed its ability to obviate and correct these alterations. Milder alterations can cause **deranged cellular function(s)** and limited or progressive failure and loss of cellular activities, e.g. progressive organ failure syndrome following overexposure to radiation. Somatic mutations affecting the proto-oncogenes or genes regulating DNA repair systems can result in **malignant transformation** of the cell and development of cancer.

Somatic cells overburdened with mutation are driven to a special suicidal pathway involving its involuntary death to protect other cells from the hazardous risk of its malignant transformation. It might also be looked at as a cellular economic adaptation behavior by getting rid of mutated diseased, energy consuming and harmful cells. This mechanism of compulsory involuntary or programmed cell death, referred to as **apoptosis**, represents an important protective anti-mutation mechanism of the organism to maintain its genomic integrity, and demonstrates one of the basic conservative features of living organisms by executing over mutated cells to prevent spread of their mutations through division to daughter generations.

Germinal mutations are mutations that affect genomes of **germ line cells**, i.e. ova and sperms, that participate in fertilization and determination of the genetic constitution of the offspring. If the particular ovum or sperm affected by the mutation happens to participate in fertilization and zygote formation, the mutation will be inherited and

transferred to all cells the offspring and a carrier or disease state might result according to the nature of the mutated gene and the pattern of its inheritance.

#### 12.5.4. Pathological versus non-pathological mutation

The deleterious effects of mutation are determined by many factors. Mutations affecting functional elements of the genome, i.e. functional genes, can result in gene dysfunction causing deficient synthesis of the gene product or synthesis of defective product. These mutations are expected to cause pathophysiological changes and disease phenotypes in affected individuals, and represent an overt example of harmful pathological disease causing mutations. On the other hand, mutations affecting non-functional regions of the genome, e.g. intergenic areas of DNA and intronic segments of genes, do not result in pathophysiological alterations and are referred to as non-pathological mutations.

Though the vast majority of mutational events of the genome are harmful, some mutations, conversely, have beneficial effects. Such mutations, known as **beneficial mutations**, exert their effects by conferring selective advantage on affected cells secondary to the resulting change of the structure and function of protein/enzyme product of the mutated gene. The most obvious example of beneficial mutations is the sickle mutation of the  $\beta$ -globin gene. Carriers, but not homozygotes, of this mutation have a selective value because they are more resistant to infestation by the malaria parasite compared to normal subjects. Another example of beneficial mutations is the deletion mutation of the human CCR5 or CD195 or chemokine receptor type 5. The protein of the normal gene acts as a T cell membrane receptor and is used by human immunodeficiency virus (HIV) as a gate to enter the cell and build up infection with the virus. Mutation of the CCR5 gene, CCR5 Delta32 deletion, results in synthesis of an altered protein that loses its receptor function and can't bind the HIV. Carriers of this beneficial mutation are resistant to infection and genetic homozygosity of the mutation is known to confer protection against infection with the virus

#### 12.5.5. Static versus dynamic mutation

Static mutations signify mutations that are transmitted **without change** in type or extent of the mutation from a parent to his progeny. For instance, a father suffering from Marfan syndrome due to a specific mutation in the fibrillin gene may transfer the mutation **as it is** to some of his offspring. Mutation analysis by molecular diagnostic techniques of the gene of both father and child will reveal identity of the mutation between them in the affected gene. Static mutations characterize the mutational events that underlie the development of most currently defined genetic diseases.

Dynamic mutations, on the other hand, represent a peculiar category of genetic mutations that increase in magnitude or extent upon transmission from a carrier parent to an offspring. The discovery of this type of mutation clarified the pathogenetic mechanisms responsible for pathogenesis of a large number of genetic diseases known as **triplet repeat expansion** disorders. The list of these diseases due

to this type of mutation is expanding progressively and includes many common and relevant genetic disorders like fragile X syndrome, Friedreich ataxia, Huntington disease, myotonic dystrophy, some types of spino-cerebellar ataxia and many others.

In this type of mutation, a normal defined number of codons, triplets of nucleotides, along a specific segment of the gene increases in number when the gene is inherited. A **threshold effect** characterizes this dynamic mutational process. Within a certain range, we can identify a normal repeat number in the normal gene, a larger repeat number in carriers or **pre-mutation** state and a much larger number in affected patients. For instance, the fragile X mental retardation gene (FMR1 gene) which is mutated in fragile X syndrome, regulates the synthesis of a protein, fragile X mental retardation protein (FMRP) which acts, partly, as RNA binding protein that regulates in a global fashion cellular protein translation and microRNA regulatory pathways, and has important functions in many organs like the nervous system and the ovaries. The **normal** transcript of the gene has a sequence consisting of six to forty-five (**6-45**) repeats of CGG nucleotides at the 5'-untranslated region. **Carrier** mothers have a larger number of these repeats ranging from **55-200** repeats. When this permutation state of the gene is transmitted to a male offspring, progressive increase, or **dynamic expansion** of the number of the repeats happens and the gene in the **affected** male offspring, or **full mutation** state, is found to contain more than **200**, sometimes much more numbers up to hundreds, of these repeats. The deleterious effects of this mutation and the resulting pathological phenotypes of the disease are due to disruption of functions of the FMR1 gene via silencing of the gene activity and cessation of synthesis of the FMRP, or through disrupting the structure of the gene leading to production of a defective protein that can't mediate its global cellular physiological regulatory functions in post-translation modifications and targeting of synthesized proteins (**Figure 26**).

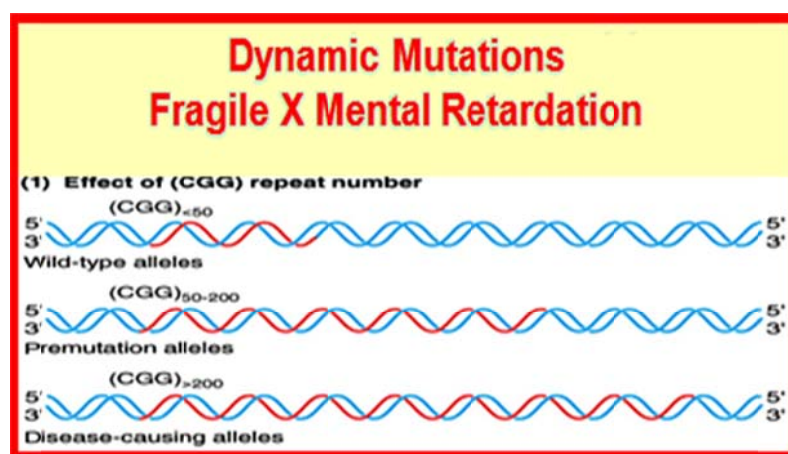


Figure 26: triplet repeat expansion in fragile X syndrome

#### 12.5.6. Persistent versus reversible mutation

Mutations are frequent structural genetic alterations in view of the dynamic nature of the genome. It almost always accompanies replication of DNA and transcription of RNA. However, most of these mutations are corrected and repaired by the genomic



repair mechanisms. Mutations in one cell that escape repair are liable to be transferred to daughter cells upon division and a whole progeny of mutated cells makes its appearance. Persistence of the mutation and its transfer to progeny cells results in **fixation** of the mutation, the change becomes persistent and appears in all daughter cells. The outcome of this scenario depends on the type of the cell and the pathogenetic effects of the resulting damage. Acquired persistent somatic mutations have the same effects as classic somatic mutations, they can cause cell death, dysfunction, apoptosis or malignant transformation. Acquired persistent germinal mutations might be transferred to offspring, become heritable and result in inherited genetic disease.

Since mutations are un-programmed genetic events, the term **reversible mutations** is actually a **misnaming** since these mutations represent **programmed structural genetic alterations** aiming at conducting specific purposes, e.g. regulation of gene function. For instance, methylation of bases to silence gene transcription, acetylation of regulatory DNA associated histones to enhance transcription and methylation of the adenine base in some types of mRNA/siRNA species to regulate energy balance or ATP homeostasis of the cell, are examples of such reversible changes of the genetic material. These alterations are sometimes considered as various mechanisms of epigenetic alterations, i.e. temporary imposed structural changes that do not alter the basic nucleotide sequence of DNA. For simplification and to obviate confusion, at least within the scope of human genetics, it might be more plausible to omit both terms, epigenetics and reversible mutations, and refer to structural genetic alterations either as mutations or as temporary structural genomic alterations.

#### **12.5.7. Point, Small, Gross, Genomic mutation**

Characterization of mutation according to extent of the resulting structural changes of the genetic material spans a very wide spectrum beginning with mutation of a single nucleotide, or even one component of the nucleotide, to mutation of the whole haploid or diploid genome. However, it is an arbitrary classification and different classification schemes can be resorted to. Classification of mutation is still a matter of debate due to lack of a standard system that can be agreed upon. Spontaneous mutations induced by transposon movement, for instance, may comprise transposition of small segments of the genetic material but can cause detrimental functional consequences on the integrity and stability of the genome. However, within this arbitrary context of classification, mutations are classified according to their extent into four main categories: point mutations, small mutations, gross mutations and genomic mutations.

##### **A. Point mutation**

Point mutation refers to mutation of one **single nucleotide**, or base, of the gene irrespective of the size of the gene or the number of nucleotides constituting the gene. Point mutations represent the commonest underlying pathogenetic mechanisms responsible for nearly 70 % of currently defined genetic disorders. Lethal genetic diseases can be caused by single point mutations, even in very large genes. A well

known example is Duchenne myopathy due to single point mutations of the dystrophin gene, which is the largest human gene composed of 2.4 million bases. Sickle cell anemia is another example of a drastic genetic disease resulting from a single point mutation of the beta globin gene. Point mutations comprise many different subtypes according to the nature of the resulting base change and the consequent alteration of gene function targeted by the mutational event (**Figure 27**).

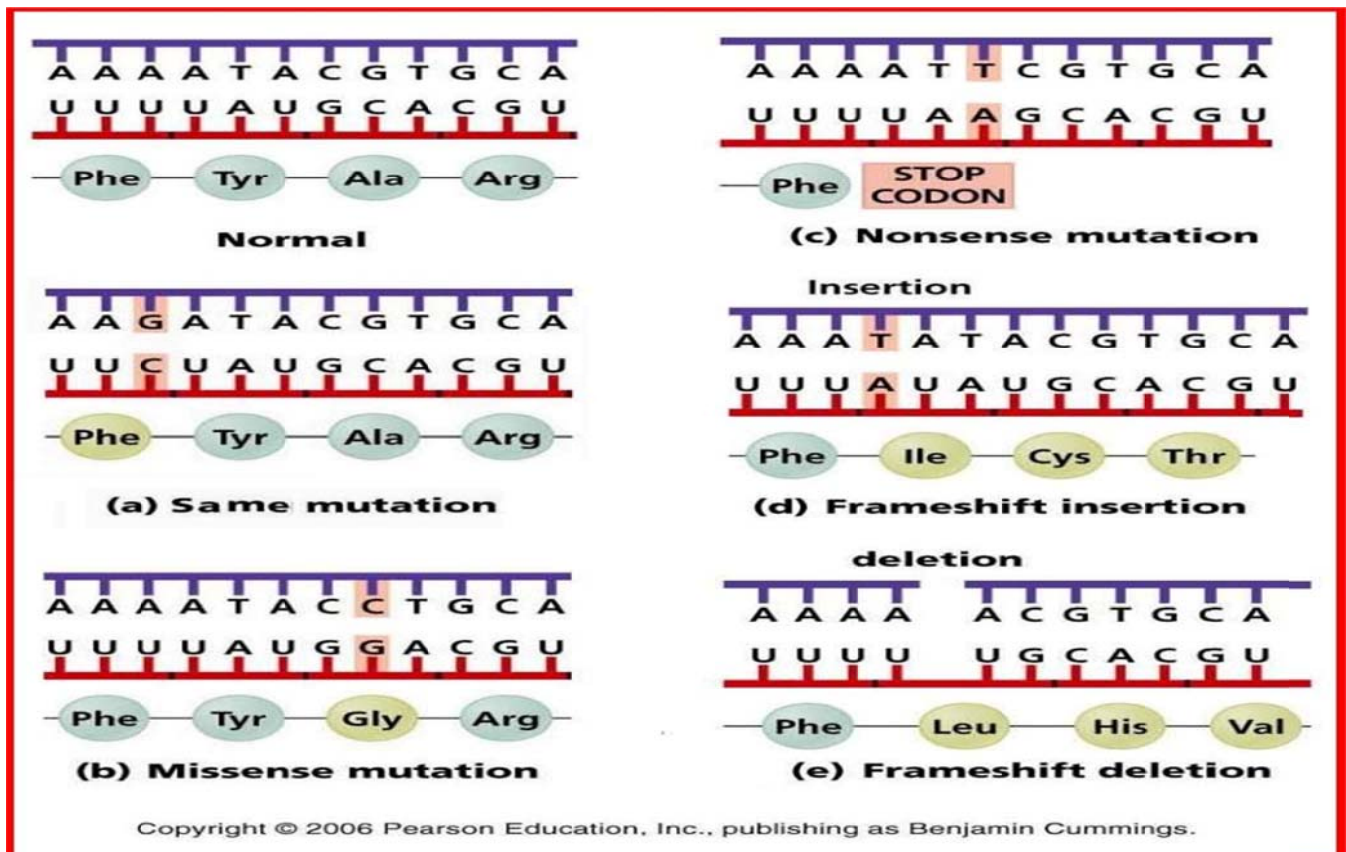


Figure 27: Types of point mutations

### Molecular mechanisms of point mutation

Molecular mechanisms of point mutation comprise a wide spectrum of alterations at the molecular level of the genetic material. These mechanisms include **replacement of one base** (nucleotide) of the gene by another base, **deletion, or loss, of one base** of the gene and **addition of one base** to the gene. Replacement of one base by another base may not affect the function of the gene or, on the other hand, can have drastic effects on gene function according to the type and consequences of the change.

If one base of a functional codon is replaced by another base changing the codon to another codon, but specifying the same amino acid, then no change in structure of the protein coded by the gene occurs, and no pathological alterations happen. This is due to degeneracy of the genetic code that allows for defining the same amino acid by different codons. This type of point mutation is called **same-sense mutation** since the

**same amino acid** is dictated by the new codon and no change in the synthesized protein results.

If the new codon specifies a **different amino acid**, a **missense mutation** results, the effects of which depend on the role played by the original amino acid in protein functioning, e.g. protein stability and aggregation. If it has important roles within fundamental structural or functional domains of the protein, that could not be performed by the new different amino acid, then pathophysiological alteration of the structure and/or function of the protein is expected and pathogenesis of a genetic defect might ensue.

If base replacement results in formation of a new **stop or termination codon**, a status termed **non-sense mutation**, the resulting alterations depend on many factors. If the mutation happens at the beginning of the coding part of the gene, most of the protein coded by the gene will not be translated from this point on. A short, or truncated, mostly non-functional protein will be synthesized and marked deterioration of gene function might result. If the mutation happens at or near the end of the coding part of the gene, most of the protein coded by the gene will be translated and less deterioration, or even no change, of gene function occurs. In both situations, the resulting pathogenetic alterations depend on the physiological role played by the missing non-translated part or domain of the protein. If these roles are important in mediating functions of vital metabolic networks in the cell, then genetic defects will happen and pathogenesis of a genetic disorder might ensue.

Deletion or addition of one base within the gene will result in shifting of the reading frame of the newly formed codons. This type of point mutation is called **frame-shift mutation**. The resulting effects may lead to formation of new same-sense, missense or termination codons, with ultimate consequences similar to the results of the aforementioned mechanisms. Frame shifting of the gene sequence leads to frame shifting in the amino acid sequence of the synthesized protein and a new protein with new sequence of amino acids might be translated. The similarity between the amino acid sequence of the wild type protein and that of the new protein depends on the start site of the frame shift mutation of the gene sequence. These same rules apply for mutational events involving the addition or deletion of two bases. If three new bases are inserted between existing codons, however, a whole new codon will be added and an additional amino acid will be added to the polypeptide chain of the protein. Also, if three bases constituting a functional codon are removed or deleted from the gene, the amino acid specified by the deleted codon will not be translated and will be missing from the synthesized polypeptide chain. Whereas addition or deletion of one codon to a functional exon of the gene might result in a wide spectrum of alterations as previously explained, whole codon addition or deletion within non-coding parts of the gene, e.g. introns, usually has no effect on translated protein or on gene function.

Change in the structure of one base of the gene, e.g. methylation of cytosine, can affect gene transcription through several different mechanisms. Each cell has a

specific **methylation pattern** necessary for normal cell differentiation during development. Changes in this pattern by **mutations of regulatory genes** controlling its timing and its magnitude can have marked deteriorating effects on the cell during differentiation. This type of mutation must not be confused with **epigenetic mutations** that involve structural changes in the chromatin or the DNA-associated proteins rather than the DNA itself.

## Types of point mutation

### 1. Mis-sense point mutations

Due to the rigid rules of the genetic code, a change of one base of a codon might turn it into another codon that defines a different amino acid in the synthesized protein. For instance, a point mutation of the codon (**TTT**) or (thymine-thymine-thymine) that defines **phenylalanine** may change it to another codon (**TCT**) or (thymine-cytosine-thymine) that defines **serine**. For protein translation system this change is considered as wrong code, hence the term missense, since it dictates a different amino acid.

### 2. Same-sense point mutations

Due to degeneracy of the genetic code, a point mutation might involve one base of a specific codon changing it to another codon, that still defines the same amino acid. For example, the amino acid **leucine** is defined or dictated by six different codons, **TTA-TTG-CTT-CTC-CTA-CTG**. Replacement of one base of some of these codons, **TTA** to **TTG**, **CTT** to **CTC** or **CTA** to **CTG**, by another base does not change the specified amino acid, leucine, in the synthesized protein, hence the term same-sense because no structural alteration of the protein synthesized by the gene happens.

### 3. Non-sense point mutations

The sixty four codons of the genetic code include three codons that do not define or dictate any amino acids and are used by the gene to signal the end of the protein synthesis process in order to form the proper protein with the proper number of amino acids. These codons (**TAA-TGA-TAG**) are termed **stop** or **termination codons** because they are recognized by both the mRNA transcript and the translation system as signals or markers of the end of the translation process since they do not define any amino acids, hence the term non-sense, so that no more amino acids are to be added to the growing polypeptide chain.

Non-sense mutations result when a functional codon defining a specific amino acid is changed to a stop or termination codon. For instance, change of cytosine in (**TAC**) codon that defines tyrosine to adenine turns it to the stop codon (**TAA**) resulting in premature cessation of translation and the production of a short or **truncated protein**. If the untranslated segment of the mRNA transcript after this pathological stop codon defines important amino acids necessary for function of the protein, functional defect results and a genetic disease might occur. Conversely, the termination or stop codon at the end of the last exon of the gene might suffer a point mutation turning it to a functional codon. In this case, part of the untranslated region of the mRNA transcript

will be included within the translation process which will continue with addition of more amino acids, not originally comprised in the protein sequence, till another stop codon is met with and recognition of the end signal and cessation of translation process occurs. The resulting abnormal pathological lengthening of the polypeptide chain results in synthesis of a **longer and larger protein**. Such abnormal polypeptides are usually unstable macromolecules that degrade easily or faster than normal correspondent chains. Also, due to their redundancy and defective primary structure, they might fail to undergo necessary post-translation modifications or post-translation targeting and trafficking to their proper cellular location. In either condition, genetic disease due to the functional deficiency of the protein might result.

#### **4. Frame shift point mutations**

Frame shift signifies change of the specific sequence of the nucleotides of the gene due to deletion or addition of nucleotide(s) followed by corresponding change of the amino acid sequence of the synthesized protein. Addition or deletion of a single nucleotide from a specific functional codon of the gene results in change of the amino acid sequence of the protein from this point up to the end of translation. Addition or deletion of two nucleotides has the same consequences on protein sequence. Addition or deletion of three nucleotides or whole codons or multiplies of them to the transcribing sequences of the gene, however, results in increasing or decreasing the number of amino acids defined by the added or deleted codons, respectively. If frame shifting mutations result in changing a functional codon to a stop codon, the results would be similar to those caused by non-sense point mutations, leading to premature cessation of translation with production of a short truncated polypeptide chain with different amino acid sequence depending on the extent of the frame shifting and the site of the mutation. Also, if frame shifting results in transforming the termination or stop codon following the last transcribing codon of the last exon of the gene to a functional codon, production of a larger and longer, usually unstable and easily degradable polypeptide chain, results with consequent deficiency of its physiological functions.

In spite of the exquisite ability of the protein translation system in recognizing and decoding the mRNA transcript and harmonizing the actions of tens to hundreds of factors involved in protein synthesis, it can't recognize changes of the original genetic code embodied within the codon sequence of the mRNA transcript. It decodes and recognizes triplets of bases, or codons, along the transcript without giving attention to whether they are comparable and complementary to the original sequences of the gene or not. It seems that the protein translation system has no **prior memory** to predict the validity of the codon sequence of the mRNA transcript in respect to both the gene sequence and the amino acid sequence of the protein. Though this defect of the translation system might be considered as a prerequisite for **evolutionary variation** of **protein phenotypes** necessary for acquisition of new functional abilities, e.g. formulation and construction of new metabolic pathways, it is a major cause of pathogenesis of genetic defects due to absence of a translation proofreading and repair system in comparison with the DNA and mRNA repair systems. This



feature is not to be confused with protein repair mechanisms conducted by **the chaperones** family of proteins which have the ability to correct some aggregation and misfolding defects of already synthesized proteins, in addition to many other important cellular functions, but have nothing to do with decoding and recognition defects during the actual process of translation and protein synthesis [Ulrich Hart et al, 2011].

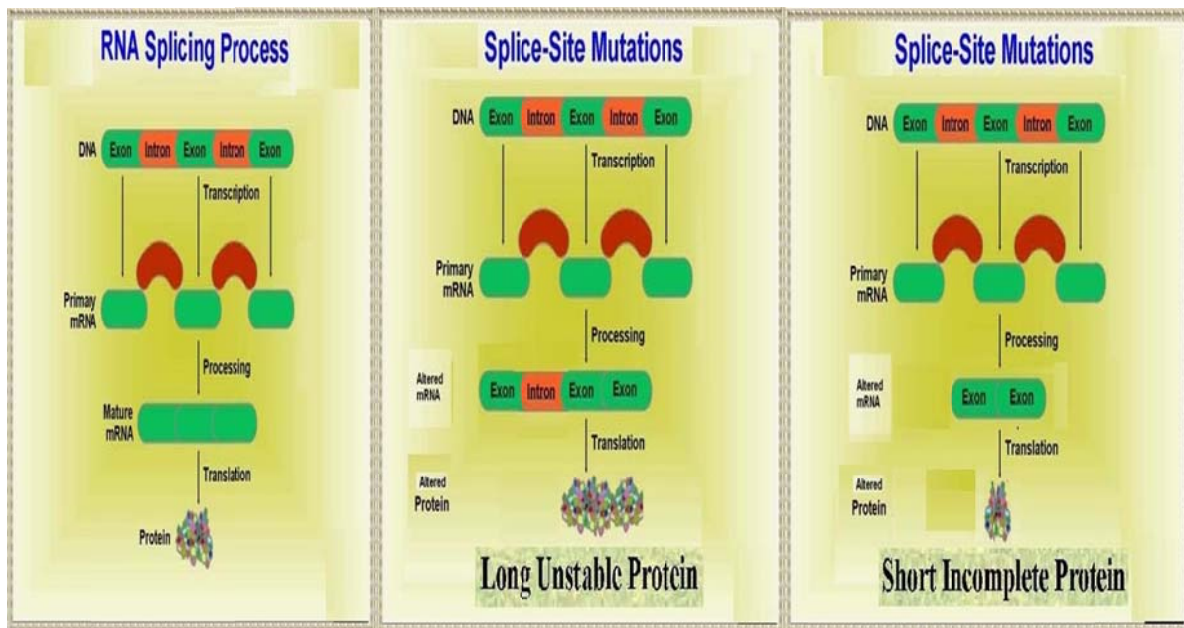
## **B. Small mutations**

Small mutations are arbitrarily classified so as to comprise structural changes involving more than one base of a gene. Accordingly, they span a very wide spectrum of pathogenetic structural alterations of the genetic material beginning with mutations involving at least two nucleotides and ending with mutations involving part of a gene, a whole gene or, even, few genes. Small mutations include **addition** or removal of two or more bases, addition or removal of one or more codons, single or multiple exon deletion or duplication, single or multiple intron deletion or duplication, **deletion** or **duplication** of a gene or of few genes, **amplification** of genes, **inactivation** of genes and **transposon-induced mutations**.

The pathophysiological consequences of these mutations depend on the resulting changes in either the **genetic regulatory networks** or the **metabolic networks** of the cell secondary to deficient and/or defective synthesis of the gene product. The resulting deterioration of gene function leading to pathogenesis of a genetic disorder correlates with the extent of deficiency of the gene product, whether due to actual deficiency of the protein or **regulatory small RNA** coded by the gene due to suppression of gene function and inability to synthesize the gene product, or to relative or absolute deficiency of the function of a faulty synthesized gene product. Deletion of large portions of a gene, multiple exon deletions for instance, result in synthesis of grossly defective protein lacking most or much of its domains needed to mediate its destined physiological roles causing marked pathophysiological deteriorations in the functioning metabolic networks dependent on these roles, thus predisposing to development of genetic disorders.

Small mutations of bases involved in splicing mechanisms mediating intron excision and splicing of remaining exons, as part of the post-transcription modifications of mRNA, can have marked effects on the translation efficiency of the final mRNA and is a well-known mutational mechanism that underlies the pathogenesis of many common and serious genetic diseases. These **splice site mutations** may result in **over excision of parts of adjacent exons** and synthesis of shorter polypeptide chains. If the missing amino acids coded by the over excised parts are important for mediating protein function or maintaining its structural integrity, then pathophysiological alterations and pathogenesis of disease might be expected. On the other hand, if splice site mutations result in **skipping of an intron**, part of an intron, a whole intron or multiple introns from excision, the resulting translated protein might be large enough to be **unstable and easily degradable**, or its structural configuration might be altered in such a way that its functional domains are

no more accessible to each other or to other components, e.g. substrates or intermediary metabolites, needed for mediating its physiological functions, and a pathogenetic defect results (**Figure 28**).



**Figure 28: Splice site mutations**

### C. Cross mutations

Gross mutations signify, primarily, **chromosomal aberrations**. In view of the distribution of the large number of human genes on the forty six chromosomes, even the smallest mutational event involving chromosomal rearrangement might comprise tens, or even hundreds, of genes. Due to the peculiar structural organization of the nuclear human genome, chromosomal abnormalities can involve either the **structure** of the chromosomes or their **number**. Mutations of chromosome structure might involve the **autosomes**, chromosomes 1-22, or the **sex chromosomes** (X and Y chromosomes). These abnormalities in structure of the chromosomes are referred to as **structural chromosomal aberrations** and include different pathogenetic types like **deletions**, **duplications**, **inversions**, **ring chromosome** formation, **unidirectional translocation**, **mutual translocations** and **Robertsonian translocations**, **chromosome gaps and breaks**, and other less common abnormal types. Chromosomal abnormalities that affect the number of chromosome are known as **numerical chromosomal abnormalities** and encompass varied group of aberrations including **trisomy**, **monosomy**, **hypodiploidy**, **heperdiploidy**, **triploidy** and **tetraploidy** (**Figures 29-A & 29-B**).

## CHROMOSOMAL MUTATIONS

Structural Mutations	Numerical Mutations
1. Deletion	1. Trisomy (47 Chromosomes)
2. Translocation	2. Monosomy (45 Chromosomes)
3. Insertion	3. Hypodiploidy (Less than 46)
4. Ring chromosome formation	4. Hyperdiploidy (More than 46)
5. Dicentric chromosome formation	5. Triploidy (3N : 69 Chromosomes)
6. Chromosome gaps and breaks	6. Tetraploidy (4N : 92 Chromosomes)

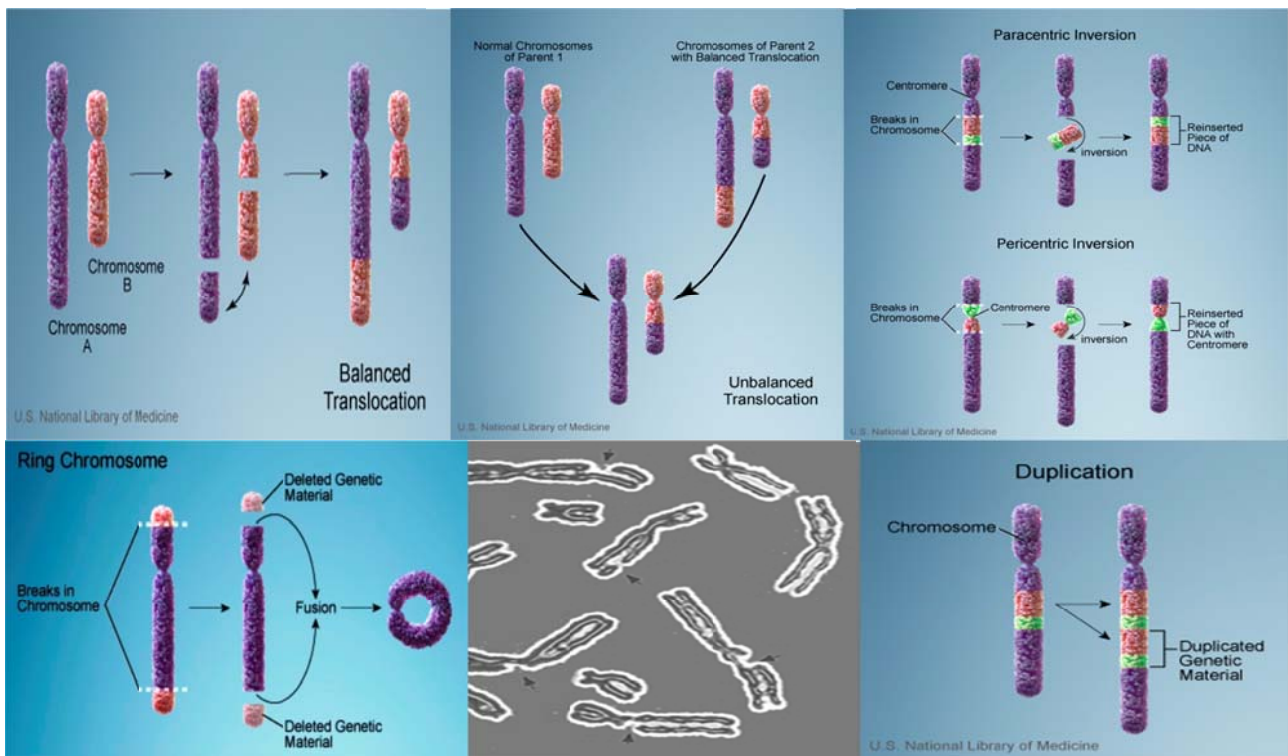
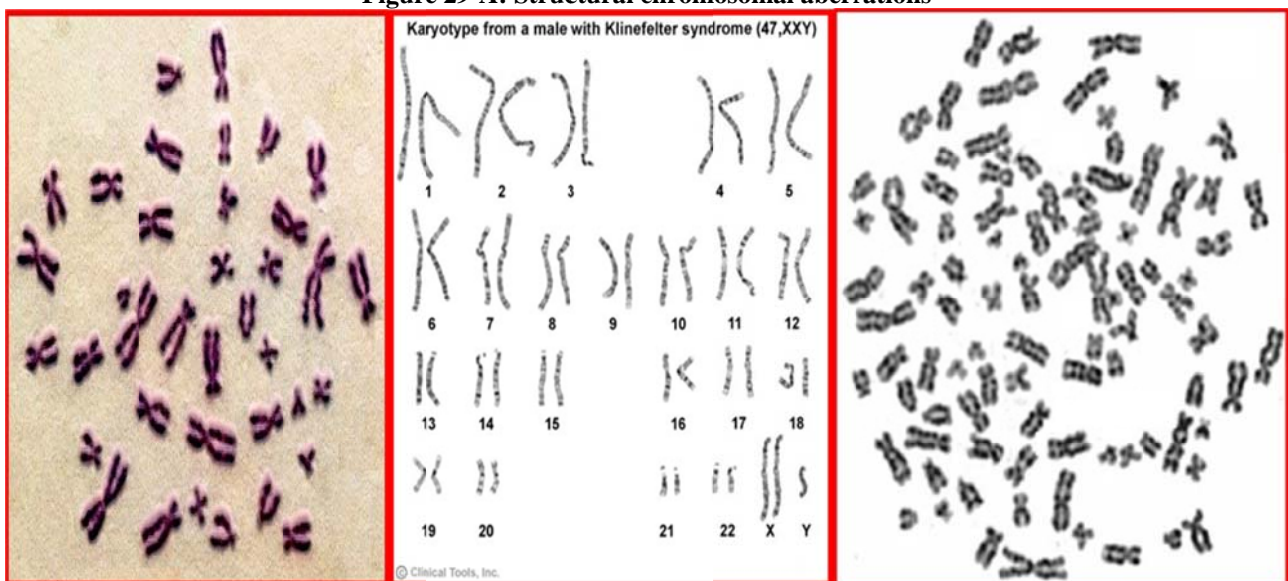


Figure 29-A: Structural chromosomal aberrations



Hypodiploidy

Hyperdiploidy

Trisomy 21 (Down syndrome)

Monosomy X (Turner syndrome)

Triploidy (69 chromosomes)

Figure 29-B: Numerical chromosomal abnormalities

## 1. Structural chromosomal aberrations

### a. Deletion

Chromosome deletion means loss of part of the chromosome. It may be **terminal** deletion when it involves the end piece of the chromosome or **interstitial** deletion when it affects other parts of the chromosome. Interstitial deletions involving the centromeric region of the chromosome lead to formation of **acentric chromosomes**. Because chromosome segregation during cell division depends on attachment of the spindle fibrils to specific centromeric receptor proteins, acentric chromosomes are lost during anaphase stage of cell division and chromosome monosomy of affected daughter cell results. Autosomal monosomy is incompatible with life of the cell in view of the deleterious functional deficiency of the proteome due to loss of hundreds, even thousands, of genes located on the missing chromosome.

### b. Translocation

Translocation involves breakage of part of the chromosome and its insertion within the genetic material of another chromosome. It may be **unidirectional** when a segment of the chromosome is broken at both ends, detached from its original site and translocated to another chromosome, or it may be **reciprocal** when two different chromosomes interchange parts of their genetic material. Translocation can result in functional deficiency and development of genetic disease if the translocated chromosome segments are inserted within intact gene sequences on the recipient chromosome leading to its disruption. Conversely, translocation might lead to over expression of the genes at the end of the translocated segment if they get inserted next to an over expressive promoter of a recipient gene. Translocations leading to over activation of oncogenes or causing disruption of critical regulatory genes, or genes responsible for maintaining genomic stability and integrity, are known to underlie the development of many malignant tumors secondary to these peculiar pathogenetic mechanisms. Examples of such cancers include Burkitt lymphoma t(8;14), follicular carcinoma of the thyroid gland t(2;3), Ewing sarcoma t(11;22), chronic myelogenous leukemia t(9;22), follicular lymphoma t(14;18) and many others.

**Robertsonian translocation** is a specific type of translocation involving translocation of a small (21-22) or large (13-15) acrocentric chromosome to another, large or small acrocentric chromosome. This defect involves fusion of the long arms of the two acrocentric chromosomes at their centromeres and loss of both short arms of the two chromosomes. Balanced carriers of this defect are phenotypically normal because they do not suffer relevant genetic deletions, since the short arms of acrocentric chromosomes do not code for essential or unique gene products.

Robertsonian translocations have an incidence of about one in a thousand newborns with its most frequent forms being between chromosomes 13 and 14, 14 and 21, and 14 and 15. Most people with Robertsonian translocations have only 45 chromosomes in each of their cells. Their children, however, might inherit the fusion chromosome and become phenotypically normal balanced carriers like their carrier parent or they may inherit a missing or an extra long arm of an acrocentric chromosome. In



unbalanced forms, Robertsonian translocations cause chromosomal deletions or addition and result in chromosome trisomy syndromes including trisomy 13 (Patau syndrome) and trisomy 21 (Down syndrome). Rarely, if both parents are heterozygous for the same translocation, they have the possibility of giving birth to homozygous viable offspring with 44 chromosomes in each of their somatic cells.

### c. Inversion

Inversion denotes breakage of a chromosome segment at both ends and its reinsertion at its original site in reversed up-down position. Inversions are classified into two main types depending on whether they include the centromere, **pericentric** inversions, or occur on one arm of the chromosome away from the centromere, **paracentric** inversions. Since inversions do not result in loss or gain of extra genetic material, inversions do not cause marked functional deficiency. However, if the break points of the inversion disrupt the sequence of an important gene, disease state might ensue. Inversions result in disturbed crossing over during meiosis and production of defective gametes with low recombination frequency. Depending on the extent of chromosome segments affected by the inversion, these defects might be incompatible with normal development and lead to early spontaneous abortions.

### d. Ring chromosome formation

Ring chromosomes are incomplete chromosome segments resulting from terminal deletion of variable parts of both chromosome arms with subsequent joining of both arms to form ring structures. The net loss of genetic material from the chromosome depends on the extent of deletion of both arms. Ring chromosome formation can also occur with terminal or total deletion of one chromosome arm only, and in rare cases it occurs when telomeres at the ends of a chromosome fuse without any loss of genetic material from the chromosome. Pathological consequences of ring chromosomes are caused by deletion of genes in the telomeric regions of affected chromosomes.

### e. Duplication

Chromosome duplication occurs when part of a chromosome is copied abnormally, resulting in extra genetic material from the duplicated segment. Duplications arise from aberrant recombination or unequal crossing-over that occurs between misaligned homologous chromosomes during meiosis with consequent reciprocal deletion of the homologous chromosomes. Chromosome duplications offer the etiological basis of gene duplication and **gene amplification**. They also offer an explanation for presence of multiple repetitive sequences along the DNA, since placement of the duplicated regions adjacent to the original sequence results in formation of **tandem repeats** along the DNA. Gene amplification due to duplication is a common phenomenon for many oncogenes because duplication and **amplification of oncogenes** results in their over expression with over production of oncoproteins that confer the malignant phenotype on transformed cells. Examples of human cancers due to oncogene duplication and amplification include esophageal cancer, gastric cancer, cervical cancer, colorectal cancer, breast cancer and neuroblastoma.

The finding that specific genomic regulatory elements, e.g. zinc finger nuclease enzyme, can induce specific chromosomal rearrangements, like duplications and inversions, suggest an evolutionary role for these chromosomal abnormalities. This event might be considered as a mechanism of **genomic evolution** through formation of novel genes that can amplify the function of parent genes or, alternatively, can undergo one or successive mutational events converting them to new genes with new functions. However, the beneficial role of this presumed **evolutionary pathway** as identified in some organisms, e.g. the mutational conversion of a duplicated digestive gene in ice fish into an antifreeze gene [Fletcher GL et al, 2001] is still a matter of debate waiting for clarification in human race since most of these chromosomal rearrangements have pathological effects on affected subjects. This might be attributed to their being an expression of **genomic instability**, rather than of genomic evolution, of the human genome.

#### f. Isochromosomes

Normally, duplicated chromosome pairs separate from each other and get equally distributed to daughter cells by the attachment of the spindle fibrils to specific receptor proteins on their common centromere with consequent disjunction along their longitudinal axis (**Figure 30**). Abnormal attachment of the spindle fibrils or instability of the centomeric region might result in abnormal separation of the chromosome pair along its transverse axis so that two chromosomes result: one consisting of the long arms of the pair and the other consisting of the short arms only, i.e. mirror-image chromosomes. This results in complete deletion and functional monosomy of the short arm or the long arm of each resulting isochromosome, respectively. Isochromosomes have equal arms with identical banding pattern of both arms. They are seen in some female patients with Turner syndrome and in many types of myeloid and lymphoblastic leukemias. They may rarely involve autosomes, e.g. isochromosome 18q syndrome.

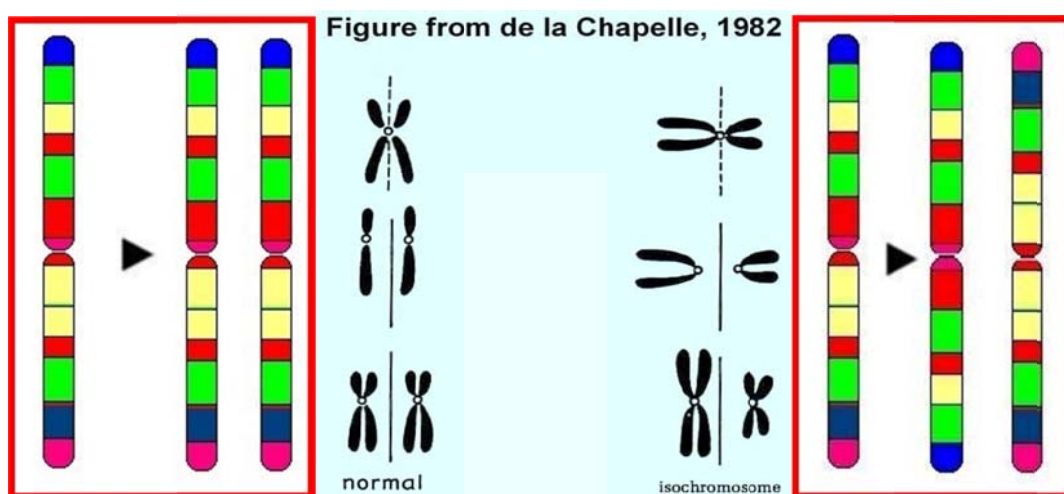


Figure 30: Isochromosome formation

### g. Mosaicism

Chromosomal mosaicism describes the presence of two or more cell populations with different chromosome complements in one individual. For instance, mosaic Down syndrome (46,XY/47,XY,+21), mosaic Turner syndrome (46,XX/45,X) and mosaic Klinefelter syndrome (46,XY/47,XXY). Normally, due to **Lyonization** or X chromosome inactivation, all female cells are **functionally mosaic** as regards the X chromosome, i.e. cells with the maternal X inactivated depend on the paternal X chromosome genes for mediating X chromosome-dependent genetic functions, and vice versa. Pathological mosaicism, on the other hand, occurs when a mutational event selectively affects one cell and get transmitted to all descendent daughter cells. It can affect gonadal cells during meiosis resulting in some gametes carrying the mutation with the rest of the germ cells being free from mutation, **gonadal mosaicism**, as well as somatic cells at any stage along the pathway of cell division after the first post-zygotic division.

Accumulating evidence suggests that chromosomal mosaicism, probably, represents a significant genetic phenomenon with widespread effects on many aspects of human biology and health since a significant proportion of human pathogenic conditions are associated with chromosomal mosaicism. Though mosaicism is a clear demonstration of **genetic diversity**, its true significance remains, as yet, unknown in view of the contradictory genetic effects attributed to and associated with its occurrence. For instance, mosaicism is quite a frequent finding among human foetuses and is observed in nearly 25% of spontaneous abortions. It also plays a role in the generation of **meiotic aneuploidy** known to be the leading genetic cause of human prenatal death, congenital malformations, prenatal mortality and postnatal morbidity, many neuropsychiatric disorders, chromosomal syndromes, learning disabilities, cancer and immune diseases. Contrariwise, there is strong and substantial evidence suggesting that increased incidence of mosaic aneuploidy in the developing human brain is an integral component of normal prenatal development of human central nervous system.

## 2. Numerical chromosomal aberrations

Numerical chromosomal abnormalities refer to chromosomal abnormalities caused by presence of abnormal number of chromosomes within the cell. These gross mutational events can involve autosomes, e.g. **trisomy** 21 in Down syndrome and **hypodiploidy** and **hyperdiploidy** in cancer cells, or the sex chromosomes like **monosomy** X (45,X) in Turner syndrome and XXY abnormality in Klinefelter syndrome (47,XXY).

**a. Trisomy**, or presence of three copies of a chromosome instead of two, results from **non-disjunction**, or non-separation, of duplicated chromosomes after replication. Normally, during cell division, replication of DNA occurs and chromosomes duplicate themselves so that each daughter cell receives the same number of chromosomes like the parent cell, in order to keep genetic identity of the species. If a duplicated chromosome pair fails to separate from each other and both chromosomes

are distributed to one daughter cell only, then trisomy of this cell results with concomitant **monosomy** of the other daughter cell. **Partial trisomy** refers to duplication of certain parts of a chromosome, e.g. the long arm or the short arm only. It has similar etiological bases like complete trisomy including non-disjunction, spindle derangements and centromeric breaks due to genomic instability.

**b. Monosomy** means presence of one single copy of the chromosome. **Autosomal** monosomy is incompatible with normal development and cells with this type of abnormalities die because of lack of the vital functions of hundreds or thousands of genes on the missing chromosome. Monosomy of the **X chromosome** is the only monosomy compatible with life in view of the peculiar nature of the genetics of the X chromosome. **Partial monosomy**, due to deletion of part of a chromosome, is a common cytogenetic abnormality seen in many genetic syndromes. The spectrum of pathological findings in each syndrome depends on the extent of functional impairment caused by lack of functions of genes lost due to the deletion.

**c. Hypodiploidy** refers to chromosome complement less than the normal diploid (46) chromosome number of somatic cells. Cells of Turner syndrome with 45 chromosomes represent the only viable hypodiploid events. Hypodiploidy is a common cytogenetic abnormality in many malignant cells. In some cases it might be so extensive that some malignant cells have chromosome number reaching near **haploidy**, or 23 chromosomes. Hypodiploidy is an important prognostic marker for the karyotypic evolution of cancer cells since the smaller the number of chromosomes a cell can live with, the more virulent malignant phenotype these cells can have. Milder states of hypodiploidy, low hypodiploidy with (33-38) chromosomes and high hypodiploidy with (42-45) are also common cytogenetic findings in malignant cells.

**d. Hyperdiploidy**, or chromosome numbers more than 46 chromosomes, underlies the pathogenesis of some genetic disorders, e.g. trisomy 13-18-21, XXY, XYY and multiple X syndromes with three or more X chromosomes in cells of affected female patients. Marked pathological hyperdiploidy with chromosome number reaching up to seventy (70) or a little bit more chromosomes, however, is quite common in many malignancies like acute lymphoblastic leukemia and acute myeloid leukemia.

### **Molecular mechanisms of chromosomal aberrations**

The pathogenetic mechanisms involved in pathogenesis of chromosomal aberrations are, still, mysterious and hardly understandable. They are not interpretable on molecular basis even for chromosome gaps and breaks in view of the complex structural assembly of the chromosomal DNA and its intimate relationship with the adjacent chromatin material. Gross **regulatory defects of mechanisms controlling cell division**, including the formation of the spindle and the timing of its action in synchronization with other biochemical and signal transduction effectors are hold responsible for predisposing to non-disjunction and the development of chromosomal **trisomies** and chromosomal **monosomies** as well. Similar defects in regulatory

mechanisms responsible for maintaining stability and integrity of the genome, thus leading to **endoreduplication**, might underlie the development of **polyploidy** conditions like triploidy and tetraploidy.

However, the regular and persistent occurrence at nearly constant incidence rates of well recognized and well defined genetic defects, including chromosomal aberrations, irrespective of ethnic, racial or environmental factors, indicates that these types of genetic changes or mutations might have, still unknown, deep rooted significance background in our genome. Disclosure of the nature of, and mechanisms of action of, master genes and related genetic factors responsible for maintaining the structural integrity, hence the identity, of the genome as a whole will, surely, throw more light on the different factors that underlie the pathogenesis and development of spontaneous chromosomal abnormalities.

#### **D. Genomic mutations**

Genomic mutations refer to mutations involving the whole genome (either the **haploid genome** of germ line cells consisting of 23 chromosomes or the **diploid genome** of somatic cells consisting of 46 chromosomes). They comprise mutations of the whole chromosome set like **triploidy** (69 chromosomes) and **tetraploidy** (92 chromosomes). They include also mutational events leading to disturbed and/or defective expression of the genome functions as a whole. **Genomic functional mutations** reveal their consequences in many phenotypes. In human, the development of vesicular or hydatidiform moles and the formation of dermoid cysts, instead of normal development of the zygote, represent obvious examples of genomic functional mutations. Normally, during the first five days following fertilization and zygote formation, all cellular processes including cell growth, differentiation and division are controlled and mediated exclusively by the maternal genome of the ovum. The sperm genome begins its participation in these processes after this critical five days post-fertilization period. **Improper temporal programming** of these genomic balances and interactions, e.g. due to **genomic imprinting mutations**, leads to **disturbed genomic regulation** of development and differentiation of the developing zygote, with drastic consequences leading to actual cessation and disturbed progress of normal development and differentiation. If the **maternal genome** fails to start functioning properly after fertilization or fails to maintain its **genetic regulation of development** over the critical immediate post-fertilization period, or if the **paternal genome** is prematurely induced to start its roles in genetic regulation of development and differentiation as a result of these genomic imprinting mutations, normal embryogenesis and/or fetal growth is brought to a stand-still.

Genomic functional mutations might, also, be caused by other pathogenetic mechanisms, one of these abnormal mechanisms entails exclusive predominance of one parental genome in zygote formation and development. The diploid genome of the zygote may be formed from two maternal haploid genomes (46, XX) without any sperm genome, or from two paternal haploid sperm genomes (46, YY) without participation of any maternal genome. These abnormalities in chromosomal



constitution of the zygote can be caused by many pathogenetic mutational events. **Endoreduplication** of one parental genome of the zygote with suppression, involution and disappearance of the other genome, or fertilization of the ovum by two sperm genomes with consequent disappearance of the ovum genome from the developing zygote can lead to pathogenesis of these genomic chromosomal abnormalities. A well-known wide spectrum of pathological embryonic and fetal malformations and abnormal conception products results from these genomic functional mutations

## 1. Numerical genomic mutations

**Polyploidy** denotes presence of more than two sets of haploid genomes in somatic cells which are normally diploid cells with 46 chromosomes. Polyploidy, up to **dodecaploidy** or twelve sets of haploid genomes, is specially common in flowering plants. It also occurs in some animals, such as goldfish, salmon and salamanders but is exceptionally rare in human cells. Polyploidy results from **abnormal replication** states of the genome due to deficient and/or defective functioning of genomic pathways involved in cell cycle regulation. Genomic replication errors reflect mutations affecting genomic stability and genomic integrity.

In human, genomic replication errors can lead to development of **triploid genome**, or genome with three sets of haploid genomes and comprising **69 chromosomes** and **tetraploid genome** with four sets of haploid genomes consisting of **92 chromosomes**. However, human triploidy might be caused by other mechanisms, e.g. abnormal fertilization processes, when an abnormal diploid ovum is fertilized by a normal haploid sperm, a condition known as **digyny** or **digynia**. Alternatively, it might also be caused by fertilization of a normal ovum by two sperms, a condition known as **diandry** or **dispermy**. Liver cells sometimes show tetraploid set of chromosomes but this is not pathological mutation-induced tetraploidy. It rather reflects relative temporal lag in completing cell division in relation to chromosomal duplication.

## 2. Structural genomic mutations

Fertilization entails penetration of the ovum by the sperm and gathering of both paternal **haploid genome** (23,Y/X) and maternal haploid genome (23,X) inside the fertilized ovum, now referred to as the **zygote** with a **diploid genome** (46,XY or 46,XX), to restore the normal chromosome number of somatic cells in order to begin the process of creating a new individual. Contribution of both maternal and paternal genomes is necessary for normal embryonic development in human. Embryos with only a maternally inherited genome and embryos with only a paternally inherited genome, referred to as **isoparental embryos**, fail to develop to term. These genomic structural mutations might develop, due to abnormal mutations during fertilization, and include two main types: **parthogenesis** and **androgenesis** (**Figure 31**).

Abnormal fertilization events might occur by different ways leading to aberrant development of abnormal zygotes. In parthogenesis, the sperm might fail to penetrate

the ovum which, upon duplicating its haploid genome, will produce an abnormal zygote with a diploid genome composed of two identical maternal, or ovum, genomes (46,XX). Another mechanism may involve participation of one polar body with the ovum genome in forming the abnormal zygote without any paternal, or sperm, contribution. This phenomenon is common in many animals that normally reproduce sexually. It might also be induced by certain chemical compounds or by electrical stimuli. In human, due to lack of male genome, this abnormal fertilization event can't proceed normally and terminates prematurely by forming a **teratoma**-like structure or a **dermoid** cyst consisting of abnormally organized mature tissues derived from different fetal cell types.

Androgenesis, on the other hand, entails inheritance of all the genome of the abnormal zygote from the paternal, or sperm, genome only. In approximately 80% of these abnormal events, an empty egg is fertilized by a single sperm, followed by duplication, or **endoreduplication**, of the sperm genome. The resulting abnormal zygote consists of diploid paternal genome without maternal contribution. In the remaining instances, approximately 20% of cases, an ovum is fertilized by two sperms and a **triploid zygote** with 69 chromosomes results. However, the paternal diploid genome dominates the early stages of growth and development, in spite of presence of a maternal haploid genome. In both cases, abnormal development of the abnormal zygote leads to the formation of an abnormal gestation product known as **hydatidiform mole** which might be complete or partial, with various histopathological variants and different chromosomal complements including diploid biparental, triploid and tetraploid sets of chromosomes.

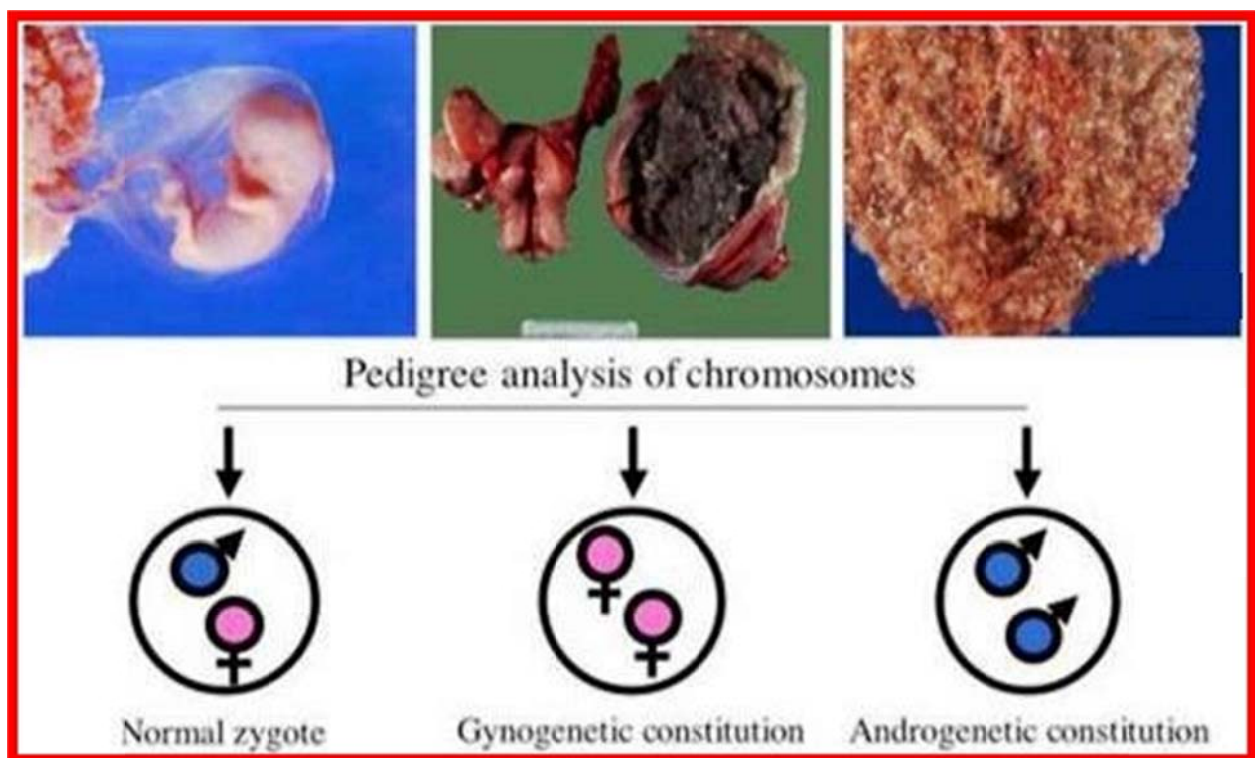


Figure 31: pathological fates of parthenogenesis and androgenesis

### 3. Functional genomic mutations

Normal development of human embryos is an extremely complex process mediated under strict control of thousands of regulatory genomic pathways. Fertilization results in bringing up two different genomes that must work together in precise concert to attain the final aim of the whole process, that is formation of a normal complete offspring. Though the immediate post-fertilization period is the most genetically active among all other periods of life of the organism, selective activation and silencing of the large number of genes constituting the genome of the zygote is a remarkable feature of this early stage. Some maternal genes are activated and others are silenced. The same process applies to parental genes as well. This phenomenon of selective activation or silencing of genes based on their parental origin is referred to as **imprinting**. Imprinting might involve a single gene, multiple genes or most genes on a chromosome as in Lyonization of the X chromosome, where it is known as gene or **genetic imprinting**, or it might involve a whole set of chromosomes of a certain parent, a condition referred to as **genomic imprinting**.

Imprinting seems to be a global biological feature of all living systems. In spite of accumulating observations, it remains to be one of the most puzzling phenomena in biology because of lack of sufficient information regarding the underlying mechanisms that mediate its effects and also due to the many species-specific and inter-species unrelated aspects and mechanisms of imprinting. For instance, regulation of expression of specific genes by imprinting is unique to therian mammals, i.e. placental mammals and marsupials, and flowering plants. The majority of imprinted genes in mammals have fundamental roles in **embryogenesis** and development of the placenta. Other imprinted mammalian genes are involved in post-natal developmental processes. Within a different biological context, for example in insects, imprinting plays a crucial role in **sex determination** where genomic imprinting of the whole paternal genome with consequent functional haploidy is a prerequisite for normal development into males.

Though many hypotheses have been proposed to explain the biological aims and the evolutionary origins of imprinting, e.g. natural selection hypothesis and parental conflict hypothesis, imprinting in essence represents a **programmed regulatory mechanism** of the genome to harmonize functions and interactions of the large number of genes that constitute it, so that integrity and stability of the genome, hence **genomic identity**, is maintained, preserved and conserved in spite of the continuously ever changing **proteome phenotype** and **proteome dynamics** of living systems. If genes constituting a specific genome are imprinted, either expressed or silenced, without strict regulatory mechanism(s) controlling temporal aspects of their functions, widespread random and conflicting qualitative and quantitative changes of the proteome synthesized by that genome will disturb the stability and integrity of the genome, through many different mechanisms, leading ultimately to decadence and loss of the genomic identity that characterizes living organisms.

### 12.5.8 Base, Sugar, Phosphate group mutation

Mutation can involve any constituent part or functional unit of the genetic material. Though point mutation denotes change of one single nucleotide of DNA, it traditionally refers to changes of the nitrogenous base of the nucleotide. Changes of the other constituents of the nucleotide, the sugar and the phosphate group, also occur and can have deleterious effects on the genetic material.

Several kinds of mutation can affect the sugar portion of the nucleotide. For instance, some physical and chemical mutagens can add oxygen to the deoxyribose of DNA or remove oxygen from the ribose of RNA, or a whole deoxyribonucleotide might be substituted by a ribonucleotide. Such mutations of DNA, containing a ribonucleotide instead of a deoxyribonucleotide are usually silent, but under abnormal conditions they might be attacked by ribonucleases leading to **DNA breaks**. Another type of mutation of the sugar portion of the nucleotide involves **methylation of ribose** of a coding nucleotide in mRNA and can lead to silencing or, even, to total suppression of translation.

Mutations that affect the phosphorus atom of the phosphate group of the nucleotide can cause widespread detrimental effects on the integrity of DNA. Particulate radiation turns ordinary phosphorus to radioactive phosphorus which, upon release of an electron, it decays into non-radioactive sulfur and suffers recoil in diameter leading to breakage of the phosphor-di-ester bonds and loss of its connection to the adjacent sugar. This results in multiple **recoil breaks** in the sugar-phosphate backbone of the DNA with ultimate breakage mutations of the nucleic acid [Irwin Herskowitz, 1977].

### 13. Pathogenesis of genetic diseases

Genetic diseases are caused by mutations. This cause and effect relationship spans a very wide spectrum of different mechanisms that follow a peculiar cascade of events beginning with mutation-induced **pathogenetic alterations** of the genetic material, leading to disturbed gene function and deficient or improper synthesis of gene products, whether these products are proteins or small regulatory RNA (miRNA) biomolecules. The ensuing defects in functioning of the regulatory and metabolic networks that control cellular activities cause multiple and variable defect-specific **pathophysiological derangements** that culminate in development and pathogenesis of genetic disease (**Figure 32**).

### **Structural Gene Mutations**



### **Regulatory Gene Mutations**

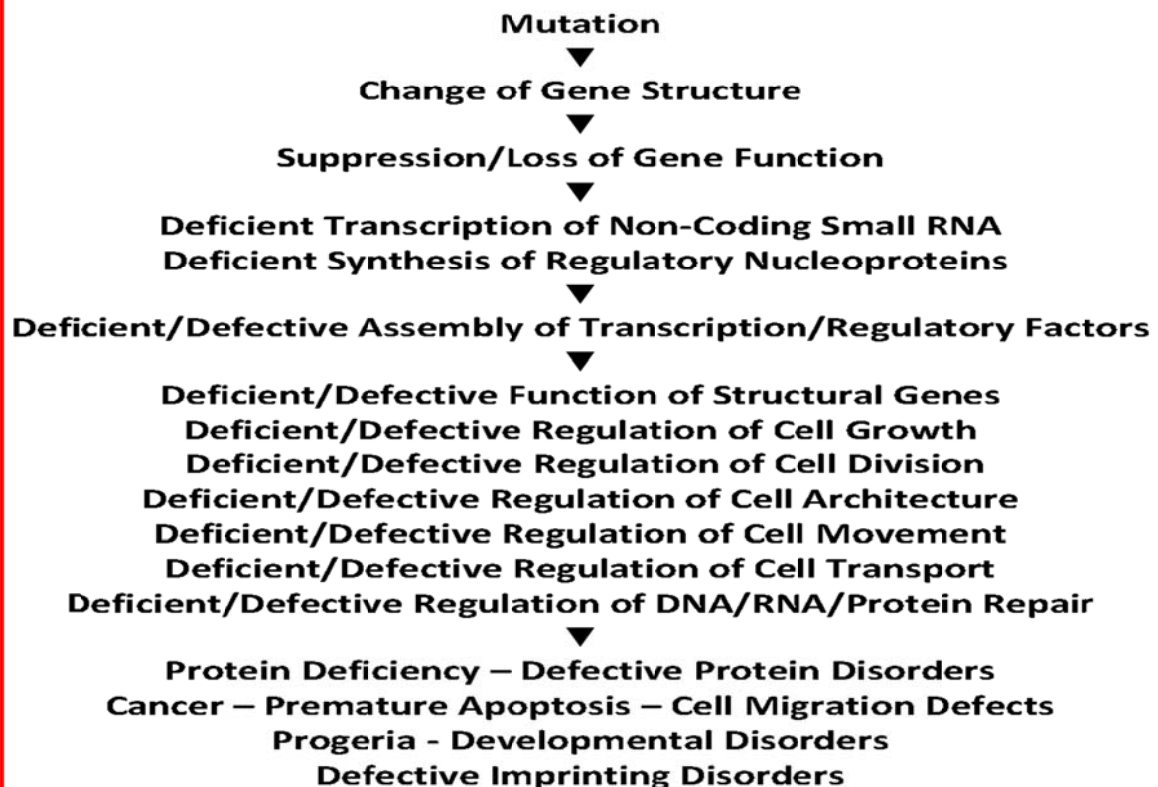


Figure 32: Pathogenesis of genetic diseases



### 13.1 The concept of protein domain and its relation to the effects of mutation

The amino acids that constitute each protein are organized, both structurally and functionally, into distinctive number of interrelated, sometimes interactive, regions called **domains**. Each domain consists of a defined number of certain amino acids arranged in a specific manner, and performs a particular role either in shaping and maintaining the structural configuration of the protein (**structural domain**), or in mediating one or more of the biological function(s) of the protein (**functional domain**). Each protein has its own specific structural configuration as regards the number, the spatial arrangement and distribution of its domains. In view of their fundamental role in maintaining structural and functional integrity and identity of the protein, protein domains are highly conserved among most species (**Figure 33**).

Mutation results in change(s) of one or more of the protein domains depending on the type of the mutation, the magnitude of its effects and the nature of the protein encoded by the mutated gene. If mutation affects one or more or all of the amino acids constituting an integral structural and/or an essential functional domain of the protein, a deleterious defect in protein function results. This pathogenetic mechanism paves the way to significant detrimental effects on the functional integrity of the protein. If the affected protein mediates specific roles in cellular activities, e.g. as a structural components of cell organelles or as a signal transducer in metabolic networks, pathophysiological alterations secondary to loss of these roles ensue with consequent pathogenesis of genetic disorder.

Alternatively, depending on the nature, the effects, the site and the magnitude of the causative pathogenetic mechanism, some mutations might primarily affect non-critical or non-integral domains of the protein. These mutations usually result in subtle conformational changes, e.g. changes in the molecular weight or the electrophoretic mobility of the protein, that do not impede, or interfere with, the physiological functions of the protein, and do not cause disease since they do not significantly disturb the **form-function relationship** that confers on each protein its biological potency.

Thus, the effects of genetic mutations are largely dependent on the resulting effects on the structural integrity as well as the functional capability of the protein. This fact clarifies, at least in part, the marked variability in potential of genetic mutations in causing disease and explains absence of any disease manifestations in spite of presence of mutation.

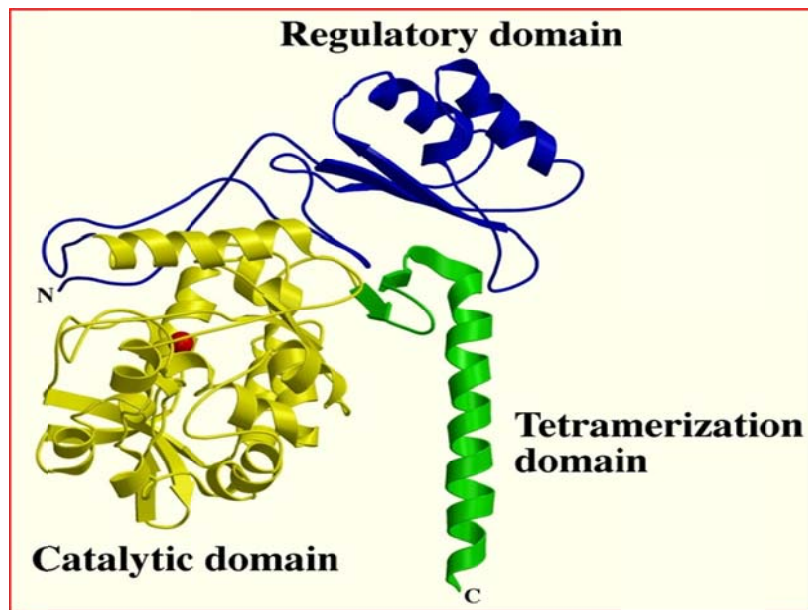


Figure 33: The concept of protein domains

### 13.2 The concept of protein mutation

The currently adopted traditional dogma of molecular biology states that the **genome**, either indirectly through the **transcriptome**, or directly in case of RNA viruses with positive strand RNA genomes, controls and regulates synthesis of the **proteome**. It also states that proteins are the actual mediators of all life activities in all living organisms. This stringent cause and effect relationship between **genome mutation** and **proteome dysfunction** holds true in most instances and represents a common final pathway underlying the development of the vast majority of currently defined genetic disorders.

The perplexing finding that some freshly, in-vitro, synthesized polypeptide chains have an inherent ability of undergoing rapid and **spontaneous folding** into its native, thermodynamically stable structural conformation, and that they can undergo some secondary configuration changes, however, raises many questions regarding its biological significance because folding of polypeptide chains is a highly collaborative and complex process involving many interactions between amino acid residues, and is attained through specific folding pathways. Accordingly, spontaneous occurrence of folding might be looked at as a **pre-programmed behavior**. The further discovery of the peculiar nature and behavior of **Prion proteins** and their confusing ability to undergo inherent **self-replication** and propagation through inducing misfolding of normal properly folded proteins in infected cells, thus transforming them to misfolded prion-like proteins, raises more suspicions regarding the inherent ability of proteins to induce specific biological alterations in nucleic acids-free systems. No other biomolecules have this innate ability of self-replication or have the potential of independent acquisition of advanced or complex structural configurational changes.

The aforementioned findings elicit that **proteins**, probably, constitute a **separate biological system** with an inherent ability of **self-organization** and **self-replication**. Additionally, they suggest that some proteins might be individually able to regulate and/or direct some aspects of life activities in living organisms. Most importantly, they raise many queries regarding the completeness of the central dogma of molecular biology and its ability to explain and interpret enigmatic problems in biology that are still waiting for disclosure. The formal similarity in structural and functional organization between genes and proteins as regards: **1.** the coded information embodied within the building units of each, nucleotides of genes versus amino acids of proteins, **2.** the peculiar internal specific organization of each, promoters/exons/introns of genes versus protein domain, **3.** the proofreading and repair systems of each, DNA proofreading and repair systems versus chaperones proofreading and repair systems, **4.** the compliance of both to rules of **natural selection** and **adaptive evolution** and **5.** their liability to encounter changes in structural organization and functional capabilities by external factors, mutagens versus protein denaturing agents, all point to a pivotal role played by proteins as fundamental biomolecules that act individually as regulators of biological systems in other, still unknown, ways independent of the well known established mechanisms dictated by the central dogma of molecular biology. The specific sequence of amino acids within the protein might signify a **protein code** responsible for determining the structure and function of the protein domain formed by certain sequences composed of certain amino acids, e.g. in a way similar to defining specific amino acids by the codon the physico-chemical properties or catalytic action of protein domains might be defined by peculiar arrangement of certain numbers of specific amino acids arranged and aggregated in a specific sequence within the domain. These assumptions which need much experimentation, however, are mere speculations based on the many biological phenomena that have not yet been explained fully by the traditional dogma of molecular biology, and on their ability to interpret some of the enigmatic aspects of these phenomena, at least in part, by reconsidering proteins as separate biological systems with inherent abilities similar to, but distinguished from, those of nucleic acids, and by revising our classic concepts of considering proteins as mere effector mediators of life activities under strict control of the genetic material.

### 13.3 Pathogenetic mechanisms

The genetic material controls life activities of the cell through regulating synthesis of proteins which directly mediate these activities. Regulatory genes, in addition, control the transcription of many classes of small RNAs that have fundamental roles in direct and feed-back regulation of most aspects of the genetic material. Mutations cause structural alterations of the genetic material. Depending on the site, nature, magnitude and effects of the mutational event as well as on the functions and importance of the mutated genes, pathogenetic mechanisms that result in **deficient synthesis of gene products**, **synthesis of defective gene products** or **disturbed regulation of cellular activities** will lead to development of genetic disorders, secondary to the ensuing pathophysiological alterations of cellular functions.

Maintaining **stability**, **integrity** and species-specific **identity** of the genome represents a prerequisite, not only for executing cell functions properly, but more fundamentally for beginning, continuation and conservation of life. It is the preservation of the collaborative and integrated intimate relationship of these three aspects of the genome that represent the real kernel and true essence of existence of all forms of life. Many pathogenetic mechanisms that drastically affect genome stability and genome integrity have been defined. These particular groups of mutations constitute important detrimental events that act via different pathways, e.g. defective genetic repair mechanisms or premature induction of apoptosis, leading ultimately to loss of genome stability and integrity and induction of cell death. On molecular or cellular level, these mutations might be considered as **life-ending mutations**. These mutations might also act in a different way through total loss of reproductive fitness and disappearance of individual genomes from the gene pool of certain species.

The spectrum of pathogenetic mechanisms and the resulting pathophysiological disturbances that underlie the development of genetic disorders is quite wide in view of the complexity of the structural organization of the genome and the strict functional specialization that characterizes each of its components. Additionally, the obscure nature and unclear functions of many components of the genetic material, undoubtedly, conceal many, still unknown, pathogenetic mechanisms and hinder proper understanding of their exact pathways. It is hoped that final completion of the human genome project might disclose the exact and complete structural organization of the human genome. However, a parallel human genome function project aiming at defining the complete functional spectrum of the genome seems to be an indispensable and imperative task in order to finalize our knowledge of our genetic material.

Currently defined pathogenetic mechanisms and pathophysiological alterations implicated in pathogenesis of genetic disorders include the following:

1. Loss/damage/duplication/inactivation of nuclear genes
2. Mutation of mitochondrial genes (mitDNA)
3. Deficient/defective DNA replication/repair
4. Triplet repeat expansion disorders
5. Loss/acquisition/damage of chromosomes
6. Deficient transcription of mRNA
7. Transcription of defective mRNA
8. Deficient/defective post-transcription mRNA repair
9. Deficient/defective post-transcription modifications of mRNA
10. Deficient translation of proteins
11. Translation of defective proteins
12. Deficient/defective post-translation modification of proteins
13. Deficient/defective post-translation repair of misfolded proteins
14. Deficient/defective post-translation targeting and trafficking of proteins

15. Deficient/defective regulation of cell growth
16. Deficient/defective regulation of cell division
17. Deficient/defective regulation of cell differentiation
18. Deficient/defective regulation of cell migration
19. Deficient/defective regulation of intercellular contact and cell movement
20. Deficient/defective apoptosis/selection repair
21. Deficient/defective regulation of cell architecture and cytoskeleton: e.g. ciliary dyskinesia disorders (bronchiectasis, dextrocardia and situs-inversus, hydronephrosis, hydrocephaly, male infertility and repeated abortions), hereditary spherocytosis, Wiskott-Aldrich syndrome and neural tube defects.
22. Imprinting disorders: genomic imprinting disorders, e.g. ovarian teratomas and hydatidiform moles, and genic imprinting diseases, e.g. Prader-Willi syndrome, Angelman syndrome, Silver-Russell syndrome, Beckwith-Wiedemann Syndrome.
23. Deficient/defective regulation of cellular functions:
  - a. Deficient/defective transport across cell membrane or membranes of cell organelles (transport defects)
  - b. Deficient/defective transport across cell pores, nuclear pores or pores of cell organelles (channelopathies)
  - c. Deficient/defective secretion of gene products (protein/enzyme deficiency disorders)
  - d. Deficient/defective excretion of metabolic waste products (storage disorders)
  - e. Deficient/defective regulation of intra and inter network reactions and interactions: signal transduction disorders: e.g. neurodegeneration, diabetes mellitus, schizophrenia and Noonan syndrome.
  - f. Deficient/defective positioning of structural proteins (cell cytoskeleton disorders)
  - g. Deficient/defective regulation of intracellular trafficking.
  - h. Deficient/defective production of cellular energy: oxidative-phosphorylation disorders.
  - i. Ubiquitination/proteasome degradation defects: e.g. Friedreich ataxia, Huntington disease, Parkinson disease, Alzheimer disease, Angelman syndrome, motor neurone disease and immunodeficiency.
  - j. Apoptosis defects: e.g. congenital malformations, autoimmune disorders, cancer and neurodegeneration.

#### **13.4 Anti-mutation mechanisms of the human genome and human proteome**

The human genome develops, persists and works in a hostile environment full of existing, and continuously generated, mutagens. Mutational events induced by external factors, which include physical, chemical and biological mutagens, have widespread detrimental effects on the stability and integrity of the **genome** as well as on the stability and integrity of the **proteome**. Additionally, further and considerable damage of the structural organization and functional capabilities of both the genome and the proteome regularly occurs on continuous and progressive basis due to the continuously generated burden of internal mutagens that result from the diverse metabolic activities of the exceedingly large number of metabolic networks of the cell. Unless a powerful and effective protective and repair system actively



participates in protecting the genome and proteome of the cell against the deleterious effects of mutations, and in efficient repair of resulting damage, maintaining the stability and integrity of both of these bio-systems that constitute the framework of life activities within the cell would have been impossible.

The human genome is endowed with a spectacular multifaceted strong anti-mutation system responsible for maintaining stability and integrity of the genome and preserving its identity. It acts by protecting the genome from the detrimental effects of mutation and by repairing mutation-induced damage. Obviously, the balance between the pathological effects of mutation and the ability of the anti-mutation system to counteract and to reduce the consequences of these effects represents the main factor that determines the likelihood of having a mutation-induced genetic disease. The human anti-mutation system comprises both **innate** mechanisms common to, and shared by, all individuals, e.g. degeneracy of the genetic code, and **acquired** aspects determined by the inherited genetic background of each human being, e.g. DNA repair system.

The human **transcriptome**, being subjected to the same mutational events that can affect, alter and damage the DNA, seems to have efficient anti-mutation mechanisms to guard against occurrence of errors during RNA transcription and to correct and repair some post-transcription defects of mRNA that can cause errors during protein translation. A separate **RNA-proofreading system** seems to exist and it probably acts during transcription by relying on the sequence complementarity information or database stored within the complementary silent or non-transcribing strand of DNA. Depending on the sequence of the active strand to ensure accurate transcription might result in improper transcription if mismatch errors occur due to, e.g. polymerase dysfunction. This assumption might, partly, explain the still un-understandable behavior of gene function which involves, seemingly needless, indirect and energy consuming mechanisms by transcribing a complementary mRNA molecule, rather than an identical mRNA, that has to be decoded again by rRNA and tRNA in the ribosome during translation.

**Table 9: Anti-mutation mechanisms of the human genome**

Anti-mutation mechanisms of the human genome	
Mechanism	Types & pathways & comments
<b>1. Structural organization of the genome</b>	1. Nuclear genome 2. Mitochondrial genome
<b>2. Structural features of DNA</b>	Complementary strand stores genetic information
<b>3. Degeneracy of the genetic code</b>	Multiple point mutations might occur without affecting synthesized protein
<b>4. Nuclear localization of DNA</b>	Physical protection of nuclear genome
<b>5. DNA-associated proteins</b>	1. Physical barriers 2. Biochemical buffers 3. Deactivating biomolecules 4. Modulation of charge transport 5. Limitation of DNA helix distortion
<b>6. Replication proofreading system</b>	Prophylactic pathway during DNA replication
<b>7. Genetic repair systems</b>	<b>A. Nuclear DNA repair</b> 1. Base excision repair 2. Nucleotide excision repair 3. Direct reversal repair 4. Mismatch repair 5. Recombination repair <b>B. RNA repair/editing system</b> Post-transcription repair/editing of some mRNA defects via guide RNA (gRNA)/editosome complex <b>C. Mitochondrial DNA (mtDNA) repair</b>
<b>8. Protein repair systems</b>	Correction of post-translation protein misfolding/aggregation by chaperones
<b>9. Silencing of transposons by piwiRNA</b>	Reduces transposon-induced mutations during development
<b>10. Antioxidant enzyme systems</b>	
<b>11. Apoptosis</b>	Prophylactic pathway against spread of mutations of heavily or lethally mutated genomes to daughter cells
<b>12. Melatonin</b>	Anti-clastogenic, anti-mutagenic, anti-carcinogenic and anti-oxidant compound.

### 13.4.1 Structural organization of the human genome

The peculiar structural organization of the human genome represents the first innate anti-mutation mechanism in view of the presence of large interspersed portions of **non-functional** intragenic, **introns**, and **inter-genic DNA sequences** and **segments** that can be mutated without having appreciable deleterious functional effects. In addition to **functional sequences** needed for synthesis of protein and of regulatory small RNA species, the human genome have a considerable amount of repetitive DNA sequences, including both **noncoding repetitive DNA sequences** and **multiple copy genes** and **gene fragments**, a large number (19000-21000) of **pseudogenes**, a considerable sizable portion (about 1/6<sup>th</sup> of the total genome size) as **pyknons**, a quite large portion (nearly 40 % of the total genome size) as **transposons** and large numbers of multiple copies of functional genes that share the same regulatory function and whose suppression or damage by mutation can be tolerated by other

genes having the same function. These peculiar structural features of the human genome allows for occurrence of mutational events in many segments of the genome without having appreciable functional defects. Even if some of these genes or DNA sequences have important roles in genome function, their presence in multiple repetitive copies can greatly reduce, or even nullify, the consequences of mutational damage resulting from affecting many copies.

The presence of multiple copies, hundreds to thousands, of mitochondrial genes within the mitochondria of each cell is crucial in obviating devastating mutation-induced damage to these vital organelles in view of their role in production of ATP. This feature of mitochondrial genome allows for considerable burden of mutations to affect it before appreciable pathological consequences result. It is estimated that mutations affecting nearly 80 % of certain mitochondrial genes might occur before pathological manifestations of mitochondrial genetic diseases make their appearance due to this multiple copy feature of mtDNA.

#### **13.4.2 Structural features of DNA**

DNA exists as a double stranded structure composed of two tightly bound strands, each strand consisting of a straight sugar-phosphate backbone with opposing nitrogenous bases linked by glycosidic linkage to the sugar of one strand and by hydrogen bonds with a complementary base on the other strand. This specific structural organization of DNA serves many purposes. It stabilizes the dynamics of the molecule, permits replication and duplication of the genetic material, protects the interiorly located bases and, most important, stores a template or copy of the genetic information ready for use in case of damage of the other strand. If small or gross mutational events affect important functional portions of the genetic material, repair mechanisms can restore the exact sequence of the damaged or lost or deleted parts through restoration mechanisms based on the complementary information of the other strand. Mutations leading to damage of corresponding segments of both strands represent a catastrophic event to the genome due to absence of the **sequence database** needed for the repair mechanism to define the exact base sequence of the newly synthesized segment in place of the deleted or grossly damaged segment.

#### **13.4.3 Degeneracy of the genetic code**

Degeneracy of the genetic code represents the third innate anti-mutation mechanism of the human genome. This feature permits the occurrence of **same-sense** point mutations in functional codons without changing the amino acid defined by the mutated codon. Since some amino acids, as a part of a specific protein domain, play critical roles in attaining and maintaining correct protein structure and in mediating proper protein function, point mutations leading to replacement of these essential amino acids by other amino acids, **missense** mutations, that can't perform the functions of the original amino acids might result in detrimental effects on the structural integrity and stability of the protein followed by deleterious consequences on physiological function of the protein. Hence, degeneracy of the genetic code allows for occurrence of many point mutations, the commonest type of mutational

events and the commonest cause of genetic disorders, without changing the final structure of the synthesized protein, thus protecting against, and obviating, the pathological effects of these mutations.

#### **13.4.4 Nuclear localization of DNA**

The localization of DNA deep inside the cell nucleus represents a fourth innate anti-mutation mechanism of the human genome because it acts as a physical barrier against many mutagens that have to overcome many obstacles of cellular defense mechanisms in order to affect the nuclear genome. These defenses include the extra-cellular environment, the cell membrane, the cytoplasmic mass, the cytoplasmic enzymes and phagocytic cellular organelles and the cytoplasmic and nuclear antioxidant enzyme systems.

#### **13.4.5 DNA-associated proteins**

The DNA-associated or DNA-binding proteins, in addition to their essential roles in regulating transcriptional processes of most genes, also play fundamental roles in protecting the DNA from the damaging effects of many mutagens, in particular the free radicals that are generated during metabolic activities of the cell. They act as **physical barriers** and **biochemical buffers**, modifiers or **deactivating biomolecules** of many chemical mutagens or damaging factors that might harm the DNA. They mediate this protective role by many mechanisms including modulation of **charge transport** of oxidative agents within the DNA, limitation of DNA **helix distortion** and regulation of protein-dependent alterations in DNA base stacking [Scott Rajski and Jacqueline Barton, 2001].

#### **13.4.6 Replication proofreading system**

Preservation of genomic identity of the organism depends exclusively on accurate replication and synthesis of two identical copies of the genome during cell division, followed by transfer, or inheritance, of each copy to each daughter cell. In this manner, all cells descendent from a parent cell have nuclear genomes identical to that of the mother cell. The majority of spontaneous point mutations of the nuclear genome are prone to occur during cell division, mostly during DNA synthesis or the replication phase of the process. The replication proofreading system acts in a prophylactic way to ensure accurate insertion or addition of the proper nucleotide to the newly synthesized strand of replicating DNA. This **prophylactic function** is fundamental to reduce the rate of inevitable replication mistakes to minimum levels that could be dealt with efficiently with the DNA repair mechanisms. In spite of the impressively fast and accurate ability of the enzymes responsible for DNA synthesis, the DNA polymerases, most of them have additional proofreading ability to ensure accurate error-free DNA replication and, hence, maintaining and preserving the stability, integrity and identity of the genome during cell division, as well as during transfer of the genetic material from parents to offspring.

### 13.4.7 Genetic repair systems

Genetic repair systems responsible for correcting and repairing many different types of point and small mutations that affect the genetic material, whether induced by exogenous mutagens or occurring secondary to endogenous spontaneous alterations, comprise both nuclear DNA repair system and mitochondrial DNA repair system. Genetic function and genetic repair represent two sides of one coin. Without the persevering continuous, active and effective surveillance exerted by the genetic repair systems to detect and repair the continuously and persistently occurring mutations, maintaining stability and integrity of the genome would be an impossible task. These repair systems consist of large numbers of **enzymes, proteins and related factors** that function in complementary and collaborative mechanisms along specific pathways, with each of them having a predefined role in the repair process. For instance, if mutation causes damage of a genetic segment consisting of sequence of nucleotides, an **endonuclease** enzyme cuts both sides of the damaged segment, followed by addition of proper nucleotides instead of the damaged or deleted ones by a **polymerase** enzyme, then a **ligase** enzyme joins the ends of the newly added segment of nucleotides to the original neighboring nucleotides by forming phosphor-di-ester bonds between the phosphate and the sugar of adjacent nucleotides, thus, regaining the sugar-phosphate backbone of the DNA [Kenji Fukui et al, 2010].

#### A. Nuclear DNA repair

Nuclear DNA repair mechanisms comprise many approaches to repair mutations of DNA. These approaches include different pathways and sub-pathways according to the type, site and extent of the mutation-induced damage and also according to the stage of cell cycle affected by the mutation. They include: base excision repair (BER), nucleotide excision repair (NER), direct reversal repair, mismatch repair, transcription-coupled repair (TCR) and recombination repair (**Figure 34**).

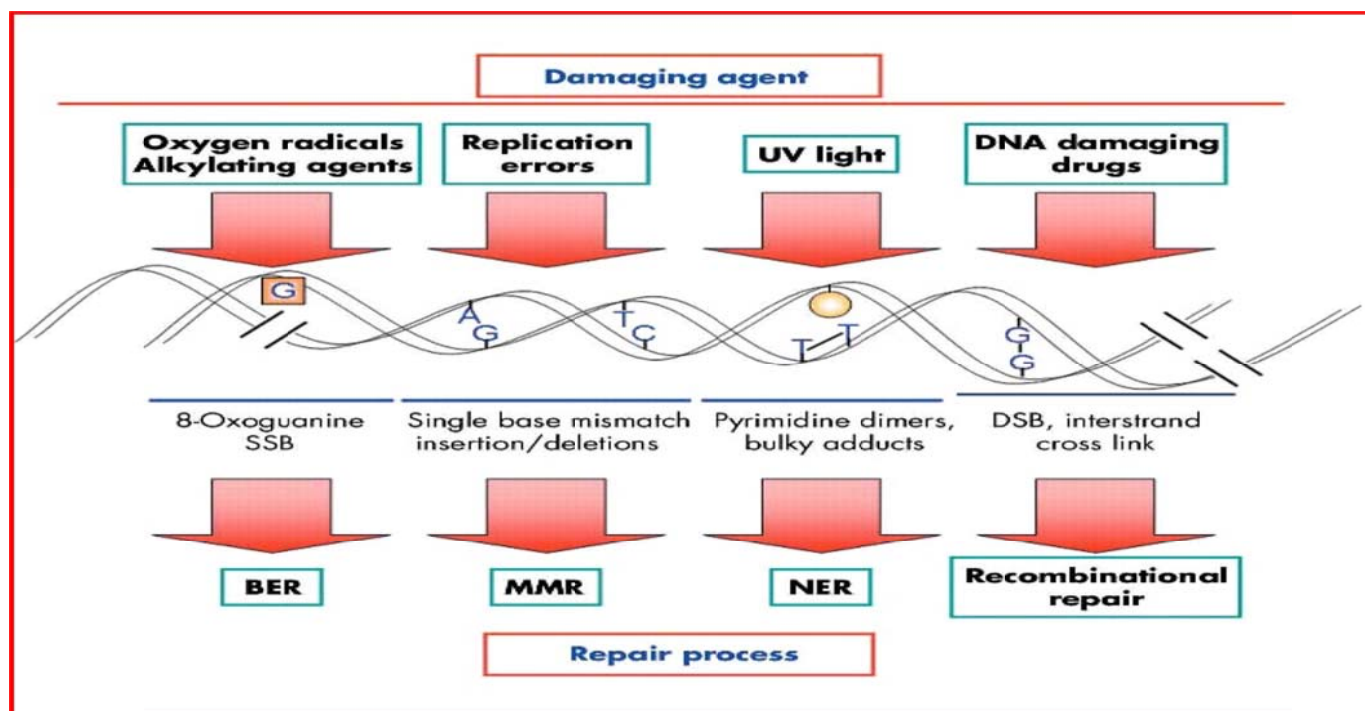


Figure 34: DNA repair mechanisms

### 1. Base excision repair (BER)

This repair mechanism is probably the most frequent DNA repair pathway in the cell. It is used for single strand point mutations affecting one or few bases of one DNA strand. It involves recognition of the damaged base of a nucleotide by a **glycosylase** enzyme and its removal by detaching it from the deoxyribose sugar via hydrolysis of the N-glycosyl bond. Breakage of the hydrogen bond between the damaged base and the opposing base on the complementary strand occurs, probably, via ATP-induced changes of the energy dynamics of the bond. Removal of the damaged or mutated base results in creation of abasic site or apurinic/apyrimidinic site (AP) of the DNA which are targeted by **endonuclease** and **lyase** activity to remove the damaged base(s) followed by addition of new normal base(s) by a specific **polymerase** enzyme and, finally, regaining the phosphodiester bonds and, hence, the phosphate-sugar backbone of the DNA strand by the action of DNA **ligase**.

### 2. Nucleotide excision repair (NER)

Nucleotide excision repair is one of the most important DNA repair systems and is highly conserved among species, though it is much more complicated in higher eukaryotes than prokaryotes. The most prominent feature of this repair system is its broad substrate specificity because it can excise DNA lesions such as UV-induced pyrimidine dimers as well as more bulky adducts of DNA.

### 3. Direct reversal repair

This repair mechanism can directly repair UV-induced pyrimidine dimer formation and alkylation adducts by DNA **photolyase** enzymes and **alkyltransferase** proteins, respectively. Direct reversal repair mechanisms are not followed by incision of DNA strands or resynthesis of new DNA since the changed or mutated bases are directly reverted to their original states either by light-dependent photoreactivation process, for pyrimidine dimers repair, or by use of alkyltransferase. Placental mammals do not have photolyase-dependent repair mechanisms and depend on nucleotide excision repair to correct and repair UV-induced pyrimidine dimer formation.

### 4. Mismatch repair (MMR)

The mismatch repair (MMR) system recognizes and corrects mismatched or unpaired bases that result from errors of DNA polymerase during DNA replication. It involves complex reactions and interactions of many enzymes, proteins and **signal discrimination factors**, probably in collaboration with the proofreading system, to recognize the mutated strand first and then to locate the site of the mismatched pair. This is followed by removal of the mutated sequence by an **endonuclease**, addition of new pair(s) by DNA **polymerase** and final regain of the DNA double stranded structure by DNA **ligases**. Post-replication mismatch repair is achieved by removal of a relatively long tract of mismatch-containing oligonucleotides, a process called long-patch MMR.



## 5. Recombination repair

Recombination repair mechanisms aim primarily at repairing double-strand breaks of DNA which represent the most devastating mutation-induced lesions of DNA because they can lead to loss of genetic information and chromosomal instabilities with consequent pathological alterations including chromosome breakage syndromes and carcinogenesis. Double-strand breaks can be caused either endogenously during DNA replication due to replication errors, e.g. replication fork collapse, or exogenously by, e.g. ionizing radiation. Recombination repair mechanisms consist of many various steps: end resection, strand invasion, DNA repair synthesis, branch migration and Holliday junction resolution, and include, at least, two different repair pathways: homologous recombination repair (HR) and nonhomologous end-joining repair. HR repair mechanism is the accurate pathway and makes use of undamaged homologous DNA as a template for repair. Non-homologous end-joining repair mechanism directly ligates two double-strand breaks ends together, and although it is efficient, it is prone to loss of genetic information at the ligation sites. However, there are many anti-recombination mechanisms to suppress excessive recombination that might lead to loss of genetic information and genomic instability [**Rihito Morita et al, 2010**].

### B. RNA repair/editing system

RNA editing refers to molecular modifications of nucleotides of RNA through chemical changes in the base makeup of the molecule. Such changes appear to be present in all three domains of life, and involve both **coding, mRNA**, and **non-coding, tRNA and mRNA**, types of small or microRNA. RNA editing occurs in the cell nucleus and the cytosol, as well as in mitochondria and is mediated by a complex repair system comprising many species of small RNA (**guide RNA**) and large protein complexes known as the **editosomes**. The pathways of RNA editing include many diverse processes: nucleoside **base modifications** such as cytidine (C) to uridine (U) and adenosine (A) to inosine (I) **deaminations**, as well as non-templated **insertions of nucleotide**. RNA editing in mRNAs effectively alters the amino acid sequence of the encoded protein so that it differs from that predicted by the genomic DNA sequence. Though mRNA editing is used in many instances to allow for synthesis by the cell of more than one protein from the same mRNA transcript, e.g. synthesis of both apolipoprotein B-100 and apolipoprotein B-48 from the same mRNA in liver cells, it can also be used to repair missense or termination mutations of the molecule which can have deleterious effects on the synthesized protein. Specific endonucleases and ligases for double stranded species of RNA have been defined in many prokaryotes and it might be just a matter of time before defining their functional counterparts in eukaryotes and human cells.

### C. Mitochondrial DNA (mtDNA) repair

The pivotal role played by the mitochondrial genome in generating ATP, without which life can neither begin nor persist, in addition to the many other critical metabolic and regulatory functions of mitochondrial genes, requires the presence of an efficient system for repairing mtDNA mutations. The need for mitochondrial

genome repair system is further imposed on the cell in view of the high mutation rate of mitochondrial genes which lack many of the anti-mutation and protective mechanisms available to nuclear genes. Similar to the nuclear genome repair system, mitochondrial repair system includes many repair pathways: base excision repair, direct reversal repair, mismatch repair, and recombination repair. Nucleotide excision repair (NER) pathway, however, seems not to be working in the mitochondria [Ricardo Gredilla, 2011].

#### 13.4.8 Protein repair systems

Accurate post-translation structural configuration of newly synthesized polypeptide chains is a fundamental conformational modification for most proteins to become functionally active biomolecules. The maturation from primary to quaternary protein structure involves many changes, e.g. folding and maintenance of steric and spatial relationships between the different domains of the protein. Conformational defects in proteins that might happen during these modifications can lead to formation of misfolded and/or aggregated non-functional molecules.

Human genome comprises a large number of genes that code a complex system composed of large numbers of specific protein families and subfamilies known as **molecular chaperones**. These proteins have many important and diverse functions in cellular activities, e.g. assisting non-covalent folding or unfolding and assembly or disassembly of macromolecular structures, including proteins. Prevention of misfolding and/or aggregation of newly synthesized polypeptide chains, which turn them to nonfunctional biomolecules, is a major and fundamental function of molecular chaperones. Other physiological functions of chaperones include: transport across mitochondrial membranes and the endoplasmic reticulum and assistance in protein degradation.

Molecular chaperones, probably, exert critical roles in maintaining stability and integrity of the proteome. This state of protein homeostasis, **proteostasis**, is a prerequisite for proper control and regulation of cellular metabolic networks by proteins and is mandatory for efficient mediation of cellular activities. Specific species of molecular chaperones, **surveillance chaperones**, are responsible for constant surveillance of the proteome to ensure proper protein homeostasis. Age-related decline or mutation-induced defects in proteome stability and integrity results in progressive aggregation and faulty conformational changes of proteins, both of which are associated with, and underlie, the development of many genetic diseases like Alzheimer disease, Parkinson disease, prion diseases and many others [Pearl and Prodromou, 2006].

#### 13.4.9 Silencing of transposon activity during development

Transposons constitute a considerable portion, nearly 40 %, of the human nuclear genome. Transposon activities might have contradictory effects on the stability and integrity of the nuclear genome. They might behave in a harmful way and act as major potential causes of **spontaneous mutations** of the nuclear genome. They can

make a copy of themselves and insert the new copy in another site, or they can detach themselves from their location and get inserted at different sites of the genome. In both conditions they result in **insertional mutagenesis** with consequent deleterious effects on genomic stability and genomic integrity. If they get inserted in a functional segment of the genome they lead to structural disruption and loss of function of the affected segment with resultant pathological effects. Alternatively, transposon activity may lead to creation and construction of new genetic combinations that may have specific functions. Within this context, they would be considered as one of the genetic biological mechanisms involved in, and responsible for, **evolutionary diversity** of the genome and the proteome. They can also cause tangible increases in the amount of the genetic material due to recurrent synthesis and addition of multiple new copies of transposable elements to the nuclear genome.

Transposons probably have a central regulatory role in early development of the embryo. A key feature of this activity involves a specific type of transposable element (MuERV-L) which is uniquely active during the totipotent cell phase of embryogenesis. A network of regulatory genes that are particularly active during early development depends on this transposable element to activate their promoter areas in order to start their transcriptional activities. The early timed enhancement of the regulatory functions of this genetic network seems essential for proper differentiation of the totipotent cells which play fundamental roles in development and differentiation. As development continues, progressive silencing of the genes of this regulatory network takes place. This delicate balance in temporal regulation of transposon activity is critical for embryogenesis and development as well as for proper development of germ line cells, particularly those in spermatocytes during spermatogenesis [Galih Kunarso et al, 2010].

Over activity or uncontrolled activity of transposons can, thus, have detrimental and devastating effects on embryogenesis, differentiation and development, and can lead to pathogenesis of a wide variety of congenital malformations and genetic defects. The human genome, however, has a unique control system composed of a specific subtype of small or micro RNA molecules, known as **piwiRNA**, or piRNA, composed of RNA-piwi protein complexes. They are thought to be involved in gene silencing, most specifically the silencing of transposons. The majority of piRNAs are **antisense to transposon sequences** suggesting that transposons are the main target of piRNA [Malone and Hannon, 2009]. In mammals, the marked activity of piRNAs in silencing of transposons and control of their activities is most important during the development of the embryo in order to reduce the rate and risk of transposon-induced mutations during this sensitive period of life.

#### **13.4.10. Antioxidant enzyme systems**

The continuous functioning of the exceedingly huge number of metabolic networks that mediate cellular activities in living cells results in continuous generation of many different types of useful and harmful metabolic by-products. Oxidant free radicals constitute one of the most crucial categories of these by-products in view of their

ability to induce widespread damage in many cellular components including membranes, organelles and structural macromolecules like lipids and proteins. This structural damage, unless counteracted by opposing antioxidant mechanisms, results in progressive degradation of cellular constituents with consequent resultant pathophysiological alterations of cellular functions, leading ultimately to disease. Although low concentrations of reactive oxygen species may be beneficial, or even necessary, in mediating many cellular processes, e.g. defense against invading micro-organisms and intracellular signaling pathways, nevertheless, higher concentrations of these free radicals play a causative role in the aging process as well as in pathogenesis of many human disease states, including immune deficiency, neurodegeneration and cancer. Oxidative damage of nDNA, mtDNA, RNA and DNA-binding proteins by free radicals represents an important category of detrimental genetic mutations induced by endogenous chemical mutagens inevitably generated during cellular metabolic activities and other cellular functions [**Valko et al, 2007**].

Living cells have several efficient non-enzymatic and enzymatic antioxidant activities that are responsible for eliminating and/or terminating the chain reactions following generation of free radicals, as a safeguard against their damaging effects on cellular components and cellular functions. Enzymatic antioxidant systems of the cell comprise large numbers of multiple types of antioxidant enzymes, notably catalase, superoxide dismutase, thioredoxin reductase, glutathione peroxidase and various other peroxidases. Efficient production of these antioxidant enzymes and proper regulation of their functions is mandatory to keep and maintain redox homeostasis of the cell which is a critical prerequisite for normal mediation of cellular activities [**Jonas and Arnter, 2001**].

#### **13.4.11 Apoptosis of heavily, lethally mutated cells.**

Apoptosis, or programmed cell death, is a universal biological behavior of most living cells necessary, in conjunction with other life-regulating mechanisms, for maintaining the vital balance between life and death that governs optimal life conditions of multicellular organisms. Apoptosis plays fundamental and crucial roles in normal growth and development as well as in normal differentiation and determination of the proper final architecture of cells, tissues and organs. Faulty timing or incorrect accomplishment of specific and selective apoptotic processes during each of these life stages of the cell might results in devastating consequences on cellular functions that range from dysfunction to malformation, and may, even, culminate ultimately in premature cell death and pathogenesis of disease [**Douglas Green, 2010**].

Apoptosis plays a crucial role in maintaining genomic stability and integrity, not of individual cells, but of the organism as a whole. Induction of apoptotic mechanisms in heavily mutated or lethally mutated cells leads to death of the cell and prevents transfer of these mutations to its putative descendant daughter cells. This fundamental **prophylactic anti-mutation role** of apoptosis in cellular activities and life prospects

of living organisms has more far-reaching effects on many important aspects related to the balance between, and the incidence of, normal and mutant genotypes within species-specific gene pools. Additionally, apoptosis can affect in an appreciable manner genomic identity of living organisms because mutation-induced evolutionary or decadence pathways are largely dependent on the outcome of certain apoptotic mechanisms operating during certain stages of the cell cycle.

#### **13.4.12 Melatonin**

Melatonin is a hormone synthesized by the pineal gland, bone marrow cells, epithelial cells and lymphocytes. Melatonin receptors are distributed in most organs, a finding reflecting its widespread roles in regulating various physiological and psychological processes. Many in vitro and animal studies revealed that melatonin has diverse functions including effective protection of cells against radiation-induced chromosome breakage [**Vijayalaxmi et al, 1995**] and inhibition of tumor development in animals exposed to experimental chemical carcinogenesis. Melatonin was shown to have protective effect against oxidative DNA damage by chemical inactivation of DNA-damaging agent as well as by stimulating DNA repair mechanisms [**Janusz Blasiak, 2007**]. These **anti-mutagenic** and **anti-clastogenic** effects of melatonin can be linked with its ability to protect DNA against oxidative damage. It may exert this **antioxidant** action by eliminating harmful reactive oxygen radicals or by stimulating the repair processes of oxidative stress-induced damage of DNA [**Tomasz Sliwinskia et al, 2007**].

## References

**Balakirev Evgeniy and Francisco Ayala** (2003): Pseudogenes: Are They Junk or Functional DNA. *Annual Review of Genetics*. 37, 123-151.

**Cheng J., Torkamani A., Peng Y., Jones T.M. and Lerner R.A.** (2012): Plasma membrane associated transcription of cytoplasmic DNA. *Proc Natl Acad Sci U S A*. 109(27): 10827-31.

**Dan Eisenberg** (2011): An evolutionary review of human telomere biology: The thrifty telomere hypothesis and notes on potential adaptive paternal effects. *American Journal of Human Biology*. 23(2): 149-167.

**Douglas R. Green** (2010): Means to an End: Apoptosis and Other Cell Death Mechanisms. Cold Spring Harbor Laboratory Press.

**Fletcher G.L., Hew C.L. and Davies P.L.** (2001): Antifreeze proteins of teleost fishes. *Ann. Rev. Physiol.*, 63, 359–90.

**Galih Kunarso, Na-Yu Chia, Justin Jeyakani, Catalina Hwang, Xinyi Lu, Yun-Shen Chan, Huck-Hui Ng and Guillaume Bourque** (2010): Transposable elements have rewired the core regulatory network of human embryonic stem cells. *Nature Genetics*. 42(7): 631-634.

**Gennadi V. Glinsky** (2009): Human genome connectivity code links disease-associated SNPs, microRNAs and pyknons. *Cell Cycle*, 8(6): 925-930.

**Gibson L.J.** (1994): Pseudogenes and Origins. *Origins*, 21(2): 91-108.

**Hatzimanikatis V., Li C., Ionita J.A. and Broadbelt L.J.** (2004): Metabolic networks: enzyme function and metabolite structure. *Curr Opin Struct Biol*, 14(3): 300-306.

**Irwin H. Herskowitz.** Principles of genetics. 2nd ed. Collier Macmillan International Editions. 1977.

**Jonas Nordberg and Elias Arnter** (2001): Reactive oxygen species, antioxidants and the mammalian thioredoxin system, *Free Radical Biology & Medicine*. 31(11): 1287-1312.

**Kenji Fukui, Noriko Nakagawa, Ryoji Masui, and Seiki Kuramitsu** (2010): Molecular Mechanisms of the Whole DNA Repair System: A Comparison of Bacterial and Eukaryotic Systems. *Journal of Nucleic Acids*, Vol. 2010, Article ID 179594.



**Malone C.D. and Hannon G.J.** (2009): Small RNAs as Guardians of the Genome. *Cell*. 136(4): 656-668.

**Mingyao Li, Isabel Wang, Yun Li, Alan Bruzel, Allison Richards, Jonathan Toung and Vivian Cheung** (2011): Widespread RNA and DNA sequence differences in the human transcriptome. *Science*, 333(6038): 53-58.

**Pearl L.H. and Prodromou C.** (2006): Structure and mechanism of the Hsp90 molecular chaperone machinery. *Annu. Rev. Biochem.* 75, 271–94.

**Ricardo Gredilla** (2011): DNA damage and base excision repair in mitochondria and their role in aging. *Journal of Aging Research*, Vol. 2011, Article ID 257093.

**Rihito Morita, Shuhei Nakane, Atsuhiko Shimada, Masao Inoue, Hitoshi Iino, Taisuke Wakamatsu, Scott R. Rajski and Jacqueline K. Barton** (2001): How Different DNA-Binding Proteins Affect Long-Range Oxidative Damage to DNA. *Biochemistry*, 40(18): 5556–5564.

**Slotkin R.K. and Martienssen R.** (2007): Transposable elements and the epigenetic regulation of the genome. *Nature Reviews Genetics*. 8, 272–285.

**Tomasz Sliwinskia, Wioletta Rozeja, Alina Morawiec-Bajdab, Zbigniew Morawiecc, Russel Reiterd and Janusz Blasiak** (2007): Protective action of melatonin against oxidative DNA damage, chemical inactivation versus base-excision repair. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*. 634(1&2): 220–227.

**Ulrich Hart, Andreas Bracher and Manajit Hayer-Hart** (2011): Molecular chaperones in protein folding and proteostasis. *Nature*, 475, 324–332.

**Valko M., Leibfritz D., Moncol J., Cronin M., Mazur M. and Telser J.** (2007): Free radicals and antioxidants in normal physiological functions and human disease. *The International Journal of Biochemistry and Cell Biology*. 39(1): 44–84.

**Vijayalaxmi A., Russel Reiterb, and Martin Meltza** (1995): Melatonin protects human blood lymphocytes from radiation-induced chromosome damage. *Mutation Research Letters*. 346(1): 23-31.

**Volkmar Weissig, Jürgen Lasch, Gregory Erdos, Helmut Meyer, Thomas Rowe and Jeffrey Hughes** (1998): DQAsomes: A novel potential drug and gene delivery system made from Dequalinium. *Pharmaceutical Research*, 15(2): 334-337.

## **Impressum / Imprint**

Bibliografische Information der Deutschen Nationalbibliothek: Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Alle in diesem Buch genannten Marken und Produktnamen unterliegen warenzeichen-, marken- oder patentrechtlichem Schutz bzw. sind Warenzeichen oder eingetragene Warenzeichen der jeweiligen Inhaber. Die Wiedergabe von Marken, Produktnamen, Gebrauchsnamen, Handelsnamen, Warenbezeichnungen u.s.w. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutzgesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Bibliographic information published by the Deutsche Nationalbibliothek: The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

Any brand names and product names mentioned in this book are subject to trademark, brand or patent protection and are trademarks or registered trademarks of their respective holders. The use of brand names, product names, common names, trade names, product descriptions etc. even without a particular marking in this works is in no way to be construed to mean that such names may be regarded as unrestricted in respect of trademark and brand protection legislation and could thus be used by anyone.

Coverbild / Cover image: [www.ingimage.com](http://www.ingimage.com)

Verlag / Publisher:

LAP LAMBERT Academic Publishing

ist ein Imprint der / is a trademark of

AV Akademikerverlag GmbH & Co. KG

Heinrich-Böcking-Str. 6-8, 66121 Saarbrücken, Deutschland / Germany

Email: [info@lap-publishing.com](mailto:info@lap-publishing.com)

Herstellung: siehe letzte Seite /

Printed at: see last page

**ISBN: 978-3-659-29835-6**

Copyright © 2012 AV Akademikerverlag GmbH & Co. KG

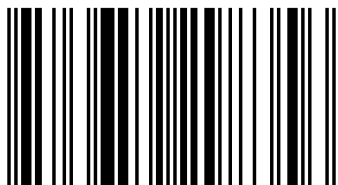
Alle Rechte vorbehalten. / All rights reserved. Saarbrücken 2012

Medical genetics, a branch of human genetics, spans a wide spectrum of sub-branches including basic, clinical, diagnostic, therapeutic, prophylactic and applied genetics. This booklet, I hope it be the first in a series, is confined to two disciplines of basic genetics, viz. molecular genetics and pathogenetics, that are concerned with studying the structure and function of the genetic material in health and disease states. It aims at offering the basic concepts of these two subjects to clinicians and research workers in the field. For clinicians beginning their specialization as medical geneticists, I tried my best to offer, as simplified as possible, definitions and illustrations of the main points of both subjects concerning the structure and function of genes, mutagens and disease-causing mutations, as well as the anti-mutation mechanisms of the human genome. For research workers, I pointed out some of the most enigmatic aspects of the field that, still, await more research to disclose in order to go a step further in our way to understand what we think we know about the role of our genetic constitution in shaping our life.



**Mohammad Salem**

Dr. Mohammad Saad Zaghloul Salem Professor of Medical Genetics Director Of Medical Genetics Center Faculty of Medicine, Ain-Shams University Cairo, Egypt. Member, Committee Of Genetic Engineering And Biotechnology National Specialized Councils Egypt.



978-3-659-29835-6